

SOCIALE ONGELIJKHEID EN ONGELIJKHEID OP BASIS VAN THUISTAAL INZAKE WISKUNDEPRESTATIES IN HET VLAAMSE ONDERWIJS

**Veranderingen tussen 2003 en 2011 op
basis van TIMSS, vierde leerjaar**

Kim Bellens, Thomas Arkens, Jan Van Damme & Sarah
Gielen
Centrum voor Onderwijseffectiviteit en -evaluatie,
KU Leuven

SOCIALE ONGELIJKHEID EN ONGELIJKHEID OP BASIS VAN THUISTAAL INZAKE WISKUNDEPRESTATIES IN HET VLAAMSE ONDERWIJS

Veranderingen tussen 2003 en 2011 op basis van
TIMSS, vierde leerjaar

Kim Bellens, Thomas Arkens, Jan Van Damme & Sarah Gielen

Promotor: Sarah Gielen

Research paper SSL/2013.03/1.1.2

Leuven, 2013



Het Steunpunt Studie- en Schoolloopbanen is een samenwerkingsverband van KU Leuven, UGent, VUB, Lessius Hogeschool en HUB.

Gelieve naar deze publicatie te verwijzen als volgt:

Bellens, K., Arkens, T., Van Damme, J., & Gielen, S. (2013). *Sociale ongelijkheid en ongelijkheid op basis van thuistaal inzake wiskundeprestaties in het Vlaamse onderwijs: Veranderingen tussen 2003 en 2011 op basis van TIMSS, vierde leerjaar*. Leuven: Steunpunt Studie- en Schoolloopbanen, rapport nr. SSL/2013.03/1.1.2.

Voor meer informatie over deze publicatie: Kim.Bellens@ppw.kuleuven.be

Deze publicatie kwam tot stand met de steun van de Vlaamse Gemeenschap, Programma Steunpunten voor Beleidsrelevant Onderzoek.

In deze publicatie wordt de mening van de auteurs weergegeven en niet die van de Vlaamse overheid. De Vlaamse overheid is niet aansprakelijk voor het gebruik dat kan worden gemaakt van de opgenomen gegevens.

© 2013 STEUNPUNT STUDIE- EN SCHOOLLOOPBANEN

p.a. Secretariaat Steunpunt Studie- en Schoolloopbanen
HIVA - Onderzoeksinstituut voor Arbeid en Samenleving
Parkstraat 47 bus 5300, BE 3000 Leuven

Deze publicatie is ook beschikbaar via www.steunpuntSSL.be

Inhoud

Beleidssamenvatting	ii
1. Inleiding	1
1.1. Wat is TIMSS?	1
1.2. Vlaanderen in TIMSS	1
2. Een eerste beschrijvende verkenning van de ongelijkheid op basis van socio-economische status en thuistaal in wiskundeprestaties in het Vlaamse onderwijs in 2003 en 2011	4
2.1. Veranderingen in leerlingenpubliek	4
2.2. Veranderingen in schoolcompositie	5
2.3. Samenhang tussen achtergrondkenmerken van leerlingen en wiskundeprestaties in 2003 en 2011	10
3. Ongelijkheid op basis van SES en thuistaal in wiskundeprestaties	15
3.1. Globale ongelijkheid op basis van SES en thuistaal in het Vlaamse wiskundeonderwijs	15
3.1.1. Invloed van achtergrondkenmerken en schoolcompositie op wiskundeprestaties in 2003 en 2011	16
3.1.2. Veranderingen tussen 2003 en 2011 inzake invloed van achtergrondkenmerken van leerlingen en schoolcompositie op wiskundeprestaties	27
3.2. Verschillen tussen scholen in Vlaanderen in sociale ongelijkheid en ongelijkheid op basis van thuistaal in wiskunde	34
3.2.1. Verschillen tussen scholen in de invloed van leerlingkenmerken en schoolcompositie op wiskundeprestaties in 2003 en 2011	35
3.2.2. Veranderingen tussen 2003 en 2011 in de verschillen tussen scholen met betrekking tot de invloed van achtergrondkenmerken op wiskundeprestaties	44
4. Referenties	48
Bijlagen	50
Bijlage 1. Steekproeftrekking, uiteindelijke steekproef en uitsluitingen in Vlaanderen in TIMSS 2011	51
Bijlage 2. Wiskundig zelfconcept: een vergelijking tussen 2003 en 2011	57
Bijlage 3. Verandering inzake leerlingkenmerken	67
Bijlage 4. “Plausibele waarden” als indicatoren van prestaties	72

Beleidssamenvatting

“De doelstelling van het decreet Gelijke Onderwijskansen is het uitwerken van een geïntegreerd Vlaams onderwijsbeleid dat alle kinderen en jongeren optimale kansen biedt om te leren en zich te ontwikkelen. Het gelijke onderwijskansenbeleid wil tegelijk uitsluiting, sociale scheiding, uitsluiting en discriminatie tegengaan en heeft daarom specifieke aandacht voor kinderen uit kansarme milieus” (Ministerie van Onderwijs en Vorming, 2012).

Het decreet omtrent gelijke onderwijskansen (GOK) trad in voegen in 2001. Hiermee trachtte de Vlaamse Overheid tegemoet te komen aan de kansenongelijkheid die bestaat in het Vlaamse onderwijs, door middel van een aantal maatregelen, waaronder het toekennen van extra financiering in de vorm van aanvullende lestijden aan scholen met een bepaald percentage van leerlingen die voldoen aan de gelijkekansenindicatoren (GOK-decreet, artikel VI.2). Kinderen die afkomstig zijn uit een gezin met een lage socio-economische status en kinderen die opgroeien in een gezin waarin de gangbare communicatie niet het Nederlands is, behoren tot de doelgroep van het GOK-beleid. Scholen met een groot aandeel kinderen die tot de doelgroep behoren, krijgen dan ook extra lestijden toegekend. Ondertussen bestaat deze regelgeving en de bijbehorende schoolpraktijken meer dan 10 jaar. Al zijn die schoolpraktijken in die periode ongetwijfeld reeds geëvalueerd, toch lijkt dit ons het geschikte moment om na te gaan of de ongelijkheid op basis van socio-economische status en thuistaal gewijzigd is doorheen de jaren.

Met dit rapport bieden we een fundamentele bijdrage tot het beantwoorden van de vraag in welke mate de ongelijkheid in het Vlaamse basisonderwijs gewijzigd is in de periode 2003-2011, wat een belangrijke indicatie biedt van de evolutie die het Vlaamse onderwijs heeft doorgemaakt gedurende het afgelopen decennium. De deelname van Vlaanderen aan zowel TIMSS 2003 als TIMSS 2011, maakt het mogelijk om eventuele veranderingen in o.a. ongelijkheid in kaart te brengen. Graag vatten we hier de resultaten die in dit rapport naar voren komen bondig samen en nemen we deze kritisch onder de loep met het oog op een partiële evaluatie van het GOK-beleid.

Allereerst zien we dat Vlaanderen stagneert met betrekking tot de gemiddelde wiskundeprestaties. Er bestaat een niet-significante tendens die aangeeft dat het prestatieniveau in 2011 lager ligt dan in 2003. Samen met het groot aantal leerlingen dat in TIMSS 2011 werd uitgesloten, is dit iets wat zeker op de voet moet worden opgevolgd. Met het steeds groter wordende aantal leerlingen die school-

lopen in het BLO (die formeel niet in het vierde leerjaar zitten) en die significant lagere prestaties behalen dan de leerlingen in het vierde leerjaar van het gewoon onderwijs, lijkt het erop dat we eenzelfde prestatieniveau voor wiskunde kunnen behouden doordat we de zwakkere leerlingen uitsluiten uit het onderzoek (en uit de populatie van leerlingen van het vierde leerjaar).

Aangezien er een aanzienlijk aantal landen in 2011 een hogere prestatie behalen in vergelijking met voorgaande jaren, lijkt deze tendens ons zeker een aandachtspunt om mee te nemen in de toekomst en is het belangrijk om de wiskundeprestaties van de leerlingen in het vierde leerjaar basisonderwijs ook in de toekomst in kaart te brengen, zodat eventuele (significante) dalingen tijdig kunnen aangepakt worden. Hierbij dient opgemerkt te worden dat er van deze minieme en niet-significante daling geen sprake meer is als we controleren voor een aantal leerlingkenmerken en de schoolcompositie inzake SES¹ en thuistaal². In dat geval zien we dat er zelfs een lichte, niet-significante stijging is van de prestaties in 2011 ten opzichte van 2003. Op basis hiervan kan aangenomen worden dat de gewijzigde leerlingpopulatie en voornamelijk de invloed van de schoolpopulatie op de prestaties van leerlingen zorgt voor extra moeilijkheden om goede prestaties te bekomen bij alle leerlingen, waardoor de wiskundeprestaties stagneren in Vlaanderen.

Daarnaast zien we dat in 2011 minder leerlingen een hoog tot geavanceerd prestatieniveau behaalden dan in 2003. Zoals aangegeven in Gielen, Bellens, Belfi en Van Damme (2012) lijkt ons onderwijs meer en meer een onderwijs voor de zwakkeren te worden; hoge presteerders worden onvoldoende uitgedaagd in het huidige Vlaamse basisonderwijs. Als gevolg hiervan zal de kloof tussen hoge en lage presteerders kleiner worden. Zoals aangehaald in Deel 3 van dit rapport, is een grote kloof – of een grote heterogeniteit in wiskundeprestaties – an sich geen slecht teken. Een grote heterogeniteit in leerlingprestaties kan er namelijk op wijzen dat elke leerling zijn/haar hoogst mogelijke prestatieniveau behaalt. Volgens ons is het ook een aandachtspunt om ervoor te zorgen dat de getalenteerde leerlingen voldoende uitgedaagd blijven, en dat ons onderwijs niet evolueert naar een onderwijs voor de zwakkeren (cfr. Dirk Van Damme, 2012). Wél vormt het een probleem als de prestaties van leerlingen samenhangen met kenmerken van hun thuissituatie en achtergrond, en met name dat leerlingen uit gunstigere thuissituaties significant hogere wiskundeprestaties behalen en omgekeerd. De mate waarin hiervan sprake is, werd dan ook uitvoerig bekeken in dit rapport.

¹ Omwille van de beschikbare informatie in TIMSS, wordt het aantal boeken thuis in dit rapport gebruikt als indicator van SES. Hoewel dit een erg beperkte indicator is en mogelijks de volledige omvang van het begrip socio-economische status onvoldoende omvat, vertoont deze indicator een sterke samenhang met de onderwijsuitkomsten van leerlingen en wordt deze indicator in de internationale literatuur algemeen aanvaard als een indicator van SES.

² Thuistaal wordt in dit rapport beschouwd als een indicator van etniciteit.

Hieruit blijkt dat wiskundeprestaties van leerlingen nog steeds samenhangen met de sociale achtergrond en de thuistaal van de leerling in Vlaanderen. Dit is zowel het geval in 2003 als in 2011. Het Vlaamse onderwijs is met andere woorden niet geëvolueerd naar meer gelijke kansen voor leerlingen met een lage SES en een andere thuistaal dan het Nederlands. Zowel lage SES leerlingen als anderstalige leerlingen ondervinden moeilijkheden in het behalen van hoge prestaties in wiskunde. Daarnaast zien we dat de schoolcompositie inzake thuistaal een bijkomend effect heeft (de invloed van schoolcompositie naar aantal boeken thuis is niet significant). Scholen met een hoger percentage anderstaligen staan voor bijkomende uitdagingen, aangezien leerlingen lagere wiskundeprestaties behalen in een school met een hoog percentage anderstaligen dan in een school met een laag percentage anderstaligen. Ook dit effect is ongewijzigd gebleven in 2011 ten opzichte van 2003.

Mede omwille van het feit dat schoolcompositie een bijkomend effect heeft op de wiskundeprestaties van leerlingen, vormt desegregatie een streefdoel in het kader van het gelijke onderwijskansen beleid. We zien echter een duidelijke tendens naar meer segregatie, althans wat betreft de segregatie naar thuistaal. Leerlingen met eenzelfde thuistaal worden meer en meer in dezelfde scholen geplaatst, waardoor we tussen 2003 en 2011 een duidelijke verschuiving vaststellen naar scholen die een overwegend Nederlandstalig publiek hebben, en scholen die een uitgesproken anderstalig publiek hebben. Met de bijkomende vaststelling in het achterhoofd dat er een grote toename van het percentage anderstalige leerlingen in het onderwijs in Vlaanderen is op te merken, is dit ons inziens een duidelijk signaal waar de nodige aandacht aan besteed moet worden om de gelijke kansen van anderstaligen te waarborgen in het Vlaamse onderwijs.

Verder zien we dat er zowel in 2003 als in 2011 in sommige scholen de samenhang tussen SES en thuistaal enerzijds en wiskundeprestaties anderzijds kleiner is dan in andere scholen. Meer in het bijzonder is de kloof tussen hoge en lage SES-leerlingen met betrekking tot hun prestaties het kleinst in scholen die hoge prestaties behalen bij lage SES-leerlingen. Voor thuistaal was dit enkel in 2003 het geval; er waren in 2003 scholen waarin de samenhang tussen thuistaal en wiskundeprestaties minder groot is dan in andere scholen. In 2011 zijn er geen verschillen tussen scholen in de samenhang tussen thuistaal en wiskundeprestaties (m.a.w. de samenhang tussen thuistaal en wiskundeprestaties is in 2011 in elke school even sterk). Dit kan mede verklaard worden door de duidelijke segregatie inzake thuistaal in het Vlaamse onderwijs, waardoor er nog weinig variatie inzake thuistaal is tussen leerlingen binnen één school.

Bovenstaande resultaten betekenen niet zonder meer dat maatregelen moeten genomen worden om de segregatie te verminderen. Evolueren naar desegregatie zorgt ervoor dat de schoolcompositie in elke school gelijk is, maar dit neemt niet weg dat deze schoolcompositie nog steeds een samenhang vertoont met de wiskundeprestaties. Zo zou de wijziging in het leerlingenpubliek naar een groter aandeel anderstaligen, ervoor zorgen dat de schoolcompositie in elke school minder gunstig wordt, en dit zou een weerslag hebben op de wiskundeprestaties van alle leerlingen in Vlaanderen. Wel dienen maatregelen genomen te worden die ervoor zorgen dat de schoolcompositie minder tot geen samenhang meer vertoont met de wiskundeprestaties van leerlingen.

Het GOK-beleid heeft getracht om hieraan tegemoet te komen tijdens het afgelopen decennium. We zien echter geen evolutie naar meer gelijke kansen in termen van socio-economische status en thuistaal in de periode 2003-2011 en er is bijkomend een tendens naar meer segregatie merkbaar. De kanttekening dient gemaakt te worden dat de gebruikte analysetechnieken in dit rapport correlatieel zijn. Dit laat toe om de samenhang tussen verschillende variabelen in kaart te brengen maar enige voorzichtigheid is geboden om causale uitspraken te doen over de directe invloed van het GOK-beleid op de ongelijkheden in het Vlaamse onderwijs. Zo is het mogelijk dat er andere factoren werkzaam zijn die een mogelijks gunstig effect van het GOK-beleid verhinderen en/of bemoeilijken. Bijkomend kan geargumenteed worden dat onderwijshervormingen meer tijd nodig hebben om zichtbare resultaten op te leveren in de onderwijspraktijk. Deze bemerkingen wijzigen echter niets aan het gegeven dat de ongelijkheid op basis van socio-economische status en thuistaal niet is afgenomen. Het huidige GOK-beleid maakt de belofte van meer gelijke kansen dus niet waar, wat in de praktijk betekent dat leerlingen met een lage socio-economische achtergrond en anderstalige leerlingen nog steeds beduidend lagere kansen hebben in het Vlaamse basisonderwijs in vergelijking met leerlingen met een hogere SES en Nederlandstalige leerlingen en deze ongelijke kansen zijn het afgelopen decennium niet verkleind. Dit lijkt ons dan ook het geschikte moment om het GOK-beleid kritisch te bekijken en bij te sturen waar nodig. Op deze manier kunnen we verhinderen dat de komende jaren opnieuw erg veel inspanningen – zowel financiële als menselijke – worden gedaan die niet leiden tot de resultaten die men hoopt (en vooropstelde) te bereiken. Hieromtrent is het citaat van een directeur basisonderwijs erg sprekend:

“Ik apprecieer ten volle alle inspanningen die de overheid (ook in de ruime betekenis: koepels, etc.) levert om ons beleid te vergemakkelijken. Door middel van extra middelen en uren maken zij het mogelijk dat er creatief kan omgesprongen worden om op school beleid te maken. Alleen worden deze middelen geleverd zonder ‘handleiding’. Zo krijgen we extra financiële middelen in functie van leerling-kenmerken, met de mededeling dat we ooit

zullen moeten verantwoorden waar deze middelen voor gebruikt worden. Een limitatieve lijst met mogelijkheden is er niet, op welke manier deze uitgaven moeten opgelijst worden krijgen we niet te horen... Idem voor de extra uren die nu net mijn bureau binnenvallen. Mijn lestijdenpakket wordt verhoogd met meerdere tientallen lestijden. Dit is grandioos en ik weet ze wel te gebruiken! En ik apprecieer ook dat we de vrijheid krijgen om deze zinvol in te vullen... Maar ik krijg het gevoel dat het beleid geschenkjes uitdeelt zonder richtlijnen. Het wetenschappelijk onderzoek omtrent gebruik van dergelijke extra lestijden wordt niet in bruikbare richtlijnen gegoten die via geijkte wegen de scholen kunnen bereiken.”

Dit citaat duidt – gestaafd met de resultaten uit dit rapport - ons inziens dan ook op een mogelijk knelpunt van het GOK-beleid, met name dat er onvoldoende richtlijnen werden aangereikt waarmee scholen en leerkrachten aan de slag kunnen gaan om de kansenongelijkheid aan te pakken. Elke school is hierdoor op haar eigen manier met de extra lestijden van slag gegaan, vanuit een eigen aanvoelen hoe zij best de ongelijkheden kan aanpakken. Een krachtig beleid dat niet enkel extra financiële middelen toekent, maar ook aangeeft op welke manier deze middelen besteed dienen te worden, is naar onze mening dan ook nodig om de kansenongelijkheid die nog steeds hardnekkig aanwezig is in Vlaanderen, effectief aan te pakken.

Wellicht vereist dit dat de overheid in haar beleid, naast de ongelijkheid, voornamelijk de algemene kwaliteit meer centraal stelt. Zoals uit de literatuurstudie van Bellens en De Fraine (2012) blijkt, hebben risicoleerlingen (in termen van lage SES en een andere thuistaal) geen nood aan ander onderwijs dan leerlingen met een hoge SES en leerlingen die Nederlands als thuistaal hebben. Kenmerken die effectief zijn voor hoge en middelmatige SES, blijken evenzeer effectief voor lage SES; kenmerken die effectief zijn voor Nederlandstalige leerlingen blijken evenzeer effectief voor anderstalige leerlingen. M.a.w., risicoleerlingen hebben geen nood aan ander onderwijs, maar het is voornamelijk belangrijk dat effectief onderwijs aan elke leerling wordt geboden. Op basis van gecoördineerd didactisch onderzoek en beleid inzake opleiding en bijscholing zal ervoor gezorgd moeten worden dat alle leerlingen een voor hen optimaal niveau bereiken. Bovendien moet, ons inziens, bevorderd worden dat scholen vroegtijdig vaststellen of hun schoolpraktijken inderdaad de gewenste resultaten bereiken en zo nodig gestimuleerd worden om deze (tijdig) bij te sturen.

1. Inleiding³

1.1. Wat is TIMSS?

TIMSS of 'Trends in International Mathematics and Science Study' is een studie die onder meer de leerlingprestaties voor wiskunde en wetenschappen in een bepaald leerjaar internationaal in kaart brengt. Leerlingen worden hierbij getoetst op kennis en inzicht van begrippen, denkprocessen en beheersing van vaardigheden. Op onderwijskundig vlak vormt TIMSS een unieke samenwerking tussen landen om de resultaten van hun onderwijs in kaart te brengen. Door het bevorderen van internationale samenwerking en overleg over leerprestaties wiskunde en wetenschappen, voorziet TIMSS de onderwijsministers van referentiecriteriën en van feedback over hoe de leerlingen van hun land presteren. De leerlingen worden getoetst na vier en acht jaar formele scholing. In Vlaanderen komt dit overeen met het vierde leerjaar van het lager onderwijs (internationaal wordt dit aangeduid met Grade 4) en het tweede leerjaar van de eerste graad van het secundair onderwijs (Grade 8).

Het TIMSS onderzoek laat toe om vast te stellen hoe het onderwijs het inzake wiskunde en wetenschappen na vier (of acht) leerjaren van formeel onderwijs doet in een internationale vergelijking (zie Gielen, Bellens, Belfi, & Van Damme, 2012). Sommige van de deelnemende leerlingen kunnen al één en zelfs twee jaar vertraging hebben (en andere leerlingen kunnen versneld zijn). De aanpak van TIMSS is op dit punt anders dan bij het PISA onderzoek, waarin gebruik gemaakt wordt van een steekproef van leerlingen van een bepaalde leeftijd, nl. 15 jaar, onafhankelijk van het leerjaar waarin ze zitten.

1.2. Vlaanderen in TIMSS

Het Vlaamse lager onderwijs nam voor het eerst deel aan TIMSS in 2003, en heeft in 2011 voor de tweede keer deelgenomen. Het Vlaamse secundair onderwijs heeft vroeger herhaaldelijk deelgenomen aan TIMSS: de leerlingen van het tweede leerjaar werden getoetst in 1995, 1999 en 2003, maar ons secundair onderwijs nam in 2011 niet deel. In 2007 nam Vlaanderen aan geen van beide onderzoeken deel. De volgende internationale meting heeft plaats in 2015. Dan zullen de leerlingen van het vierde leerjaar basisonderwijs uit 2011 in het tweede leerjaar secundair onderwijs

³ Gebaseerd op Gielen, Bellens, Belfi & Van Damme (2012)

zitten. Gedetailleerde informatie over de deelnemende leerlingen in Vlaanderen in TIMSS 2011 kan gevonden worden in Bijlage 1.

De gebruikte prestatieschaal in TIMSS werd ontworpen in 1995 en ze heeft een bereik van 0 tot 1000 punten, met een toenmalig gemiddelde van 500 en een standaarddeviatie (d.i. een maat voor de spreiding van de scores rond dit gemiddelde) van 100. De schaal wordt constant gehouden over de verschillende afnamemomenten zodat bv. de gemiddelde prestaties van de leerlingen van een land na een periode van 4, 8 of meer jaar met elkaar vergeleken kunnen worden. Hierdoor kunnen prestaties van landen overheen de jaren vergeleken worden.

Naast prestatietoetsen voor leerlingen worden in TIMSS ook achtergrondvragenlijsten afgenomen bij scholen, leerkrachten en leerlingen. Een aantal achtergrondkenmerken werden zowel bevraagd in 2003 als in 2011, wat ervoor zorgt dat niet enkel de veranderingen in prestaties in kaart kunnen gebracht worden, maar dat ook de veranderingen voor deze kenmerken ten opzichte van 2003 in kaart kunnen gebracht worden. In het verleden werden door Belfi, Arkens, Van Damme & Gielen (2012) een aantal analyses uitgevoerd met betrekking tot het wiskundig zelfconcept van leerlingen in 2003 en 2011. De resultaten staan samengevat in Bijlage 2.

In dit rapport spitsen we ons toe op prestaties in wiskunde en de samenhang tussen wiskundeprestaties en (1) de socio-economische status (SES) en thuistaal van leerlingen en (2) de schoolcompositie inzake SES en thuistaal (d.w.z. het percentage anderstaligen). De informatie over socio-economische status en thuistaal werd verkregen door middel van de achtergrondvragenlijsten die afgenomen werden bij de leerlingen. Omdat uit eerdere TIMSS-studies gebleken is dat het achterhalen van het opleidingsniveau van de ouders bij leerlingen van het vierde leerjaar veel ontbrekende waarden oplevert, is al sinds 1995 in de TIMSS-leerlingvragenlijst als proxy-indicator voor SES 'het aantal boeken thuis' opgenomen (Meelissen & Drent, 2008). Het aantal boeken thuis wordt dus gebruikt als indicator van de SES van leerlingen. De thuistaal van de leerling wordt gebruikt als indicator van de etniciteit van de leerlingen. De schoolcompositie met betrekking tot SES en het percentage anderstaligen op school worden berekend op basis van de geaggregeerde score van de leerlingen per school. Met name wordt per school het gemiddelde aantal boeken berekend en het percentage anderstaligen en deze scores worden gebruikt als indicatoren van de schoolcompositie inzake SES en het percentage anderstaligen in elke school. In dit rapport worden de data van TIMSS 2011 beschreven en geanalyseerd. Daarnaast biedt de deelname van Vlaanderen aan TIMSS in 2003 en in 2011 de mogelijkheid om de veranderingen van het Vlaamse onderwijs tussen beide tijdstippen in kaart te brengen. Met het tienjarige bestaan van het GOK-decreet in 2012 in het

achterhoofd, spitst dit rapport dan ook toe op een vergelijking van de ongelijkheid op basis van socio-economische status en thuistaal in het Vlaamse onderwijs tussen 2003 met 2011. Op deze manier wordt met dit rapport een partiële evaluatie van het GOK-decreet bekomen.

2. Een eerste beschrijvende verkenning van de ongelijkheid op basis van socio-economische status en thuistaal in wiskundeprestaties in het Vlaamse onderwijs in 2003 en 2011

In dit deel geven we aan de hand van een aantal beschrijvende cijfergegevens weer in welke mate er indicaties zijn van ongelijkheid op basis van socio-economische status en thuistaal in het Vlaamse onderwijs met betrekking tot wiskundeprestaties. In een eerste paragraaf geven we aan hoe het leerlingenpubliek gewijzigd is in de periode 2003-2011 (zie ook Bijlage 3). In een tweede paragraaf bekijken we de wijzigingen in schoolsamenstelling naar SES en thuistaal, waarmee we willen nagaan in welke mate er sprake is van (de)segregatie in 2003 en in 2011. Tot slot bekijken we in een laatste paragraaf in welke mate een samenhang bestaat tussen een aantal leerlingkenmerken en wiskundeprestaties. Deze zaken geven ons een aantal eerste indicaties van de mate van ongelijkheid op basis van socio-economische status en thuistaal betreffende wiskundeprestaties in het Vlaams onderwijs in 2003 en 2011.

2.1. Veranderingen in leerlingenpubliek

Tabel 1 geeft het gemiddelde van de leerlingkenmerken aantal boeken thuis en thuistaal weer, in 2003 en 2011. De t-toets toont aan dat het gemiddelde aantal boeken thuis significant (lichtjes) gedaald is. Dit is vermoedelijk grotendeels toe te schrijven aan de huidige tendens naar digitalisering, met o.a. de opkomst van de digitale boeken. Het percentage leerlingen dat thuis soms of nooit Nederlands spreekt is significant gestegen, wat erop wijst dat er in 2011 significant meer anderstalige leerlingen zijn in vergelijking met 2003. Belfi et al. (2012) bekeek in detail de verdeling van deze leerlingkenmerken, alsook de verandering van de leerlingkenmerken geslacht, geboortemaand en al dan niet normaalvorderend zijn. Deze resultaten zijn opgenomen in Bijlage 3.

Tabel 1. Leerlingenpubliek in 2011 ten opzichte van 2003, naar aantal boeken thuis en thuistaal.

	2003		2011		t-toets Gem. 2003 ≠ 2011
	Gem.	S.D.	Gem.	S.D.	
Aantal boeken thuis ^a	2.00	1.07	1.94	1.05	2.35*
Thuistaal ^b	.16	.37	.25	.43	-11.08***

* p < .05; ** p < .01; *** p < .001

^a 1 = 0-10 boeken; 2 = 11-25 boeken; 3 = 26-100 boeken; 4 = 101-200 boeken; 5 = meer dan 200 boeken

^b Percentage leerlingen dat soms of nooit Nederlands spreekt thuis

2.2. Veranderingen in schoolcompositie

Tabel 2 toont de gemiddelde schoolcompositie in 2003 en 2011 en de verschillen tussen beide afnamejaren. Op te merken is dat de gemiddelde schoolcompositie in termen van SES lichtjes, maar significant gedaald is, d.w.z. dat het schoolgemiddelde inzake het aantal boeken thuis lichtjes gedaald is. Daarnaast zien we dat het percentage anderstaligen (leerlingen die soms of nooit Nederlands spreken thuis) per school opmerkelijk en significant gestegen is in 2011 ten opzichte van 2003. Deze cijfers liggen in lijn met de vaststelling dat er op individueel niveau een daling in aantal boeken en een stijging van het aantal anderstalige leerlingen op te merken is tussen 2003 en 2011.

Tabel 2. Gemiddelde schoolcompositie SES en % anderstaligen per school in 2003 en 2011, met een ongepaarde t-toets getest of het verschil tussen 2003 en 2011 significant is.

	2003			2011			t-toets Gem. 2003 ≠ 2011
	Gem. (S.D.)	Min.	Max.	Gem. (S.D.)	Min.	Max	
Schoolcompositie SES ^a	2.00 (1.32)	.56	2.85	1.95 (1.23)	.59	2.75	7.32***
% Anderstaligen ^b	16.54 (.63)	0	86.67	24.99 (.77)	0	92.86	-29.14***

* p < .05; ** p < .01; *** p < .001

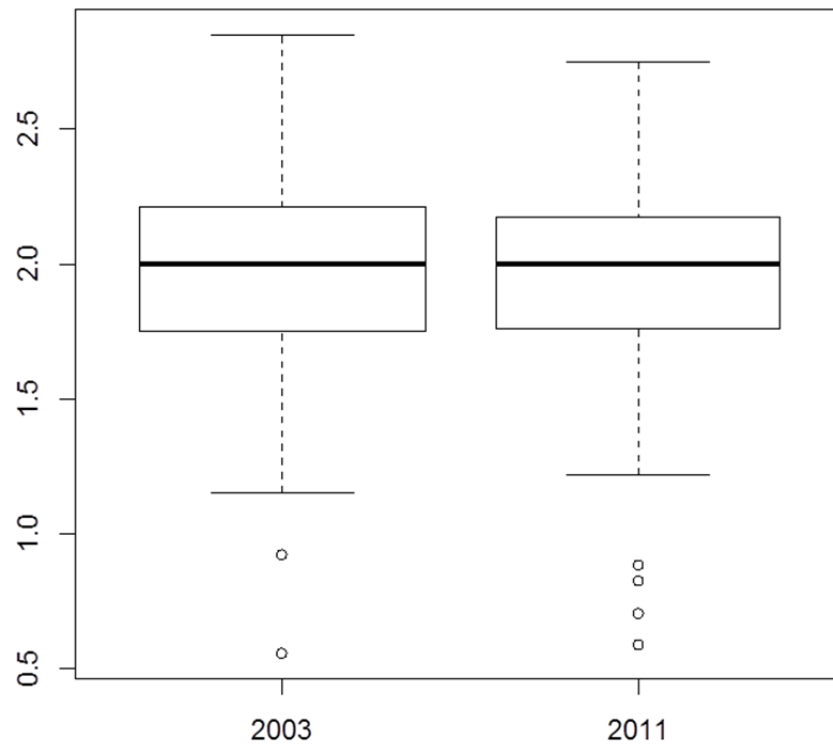
^a Geaggregeerde variabele van het aantal boeken thuis. Per school werd het gemiddelde berekend van de variabele aantal boeken thuis (voor de categorieën, zie Tabel 1).

^b Percentage anderstaligen per school.

Om een beter zicht te krijgen op de verdeling van de leerlingen over scholen en op deze manier een zicht te krijgen op de mate van (de)segregatie in het Vlaamse onderwijs, bekijken we deze gegevens omtrent schoolcompositie meer in detail. Grafiek 1 toont de spreiding tussen scholen van de gemiddelde schoolcompositie inzake SES. In deze grafiek zien we dat het schoolgemiddelde

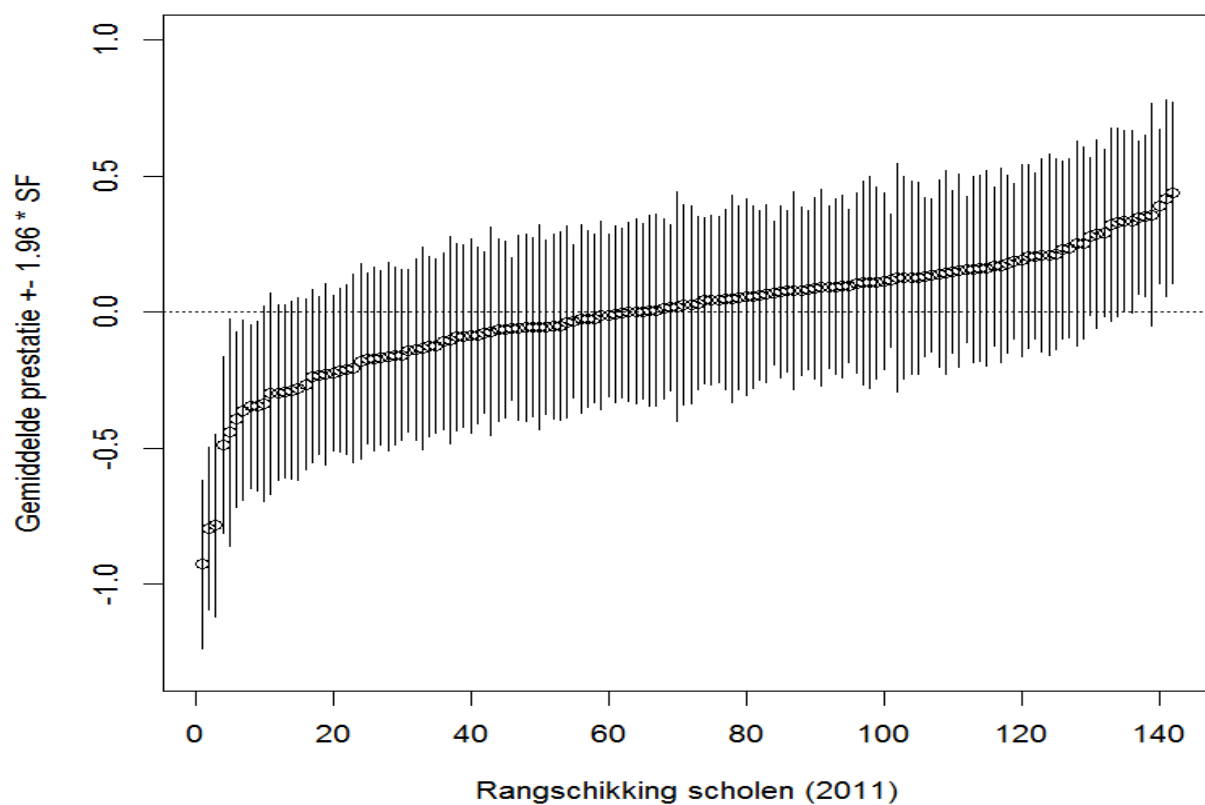
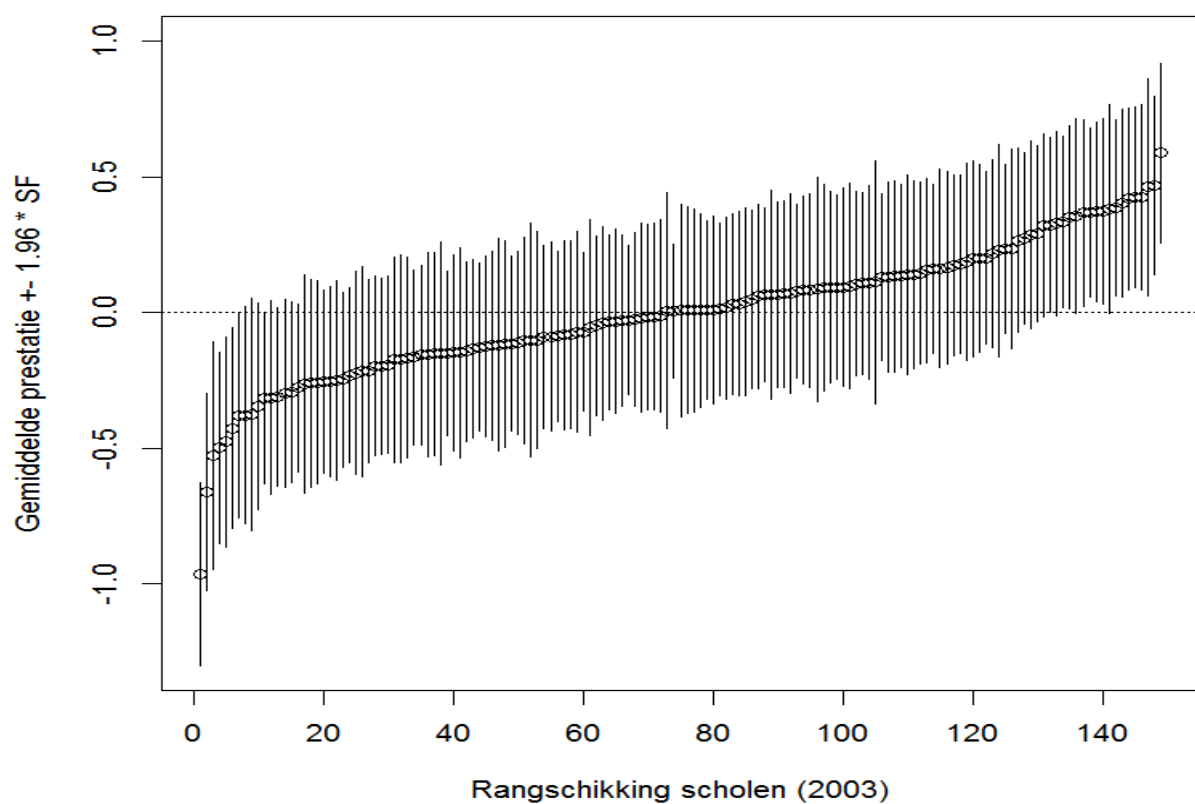
betreffende aantal boeken thuis licht gedaald is en dat dat de spreiding (iets) minder groot is in 2011 dan in 2003. Dit zijn echter erg kleine verschillen.

Grafiek 1. Boxplot verdeling gemiddelde aantal boeken thuis per school (in categorieën), in 2003 en 2011



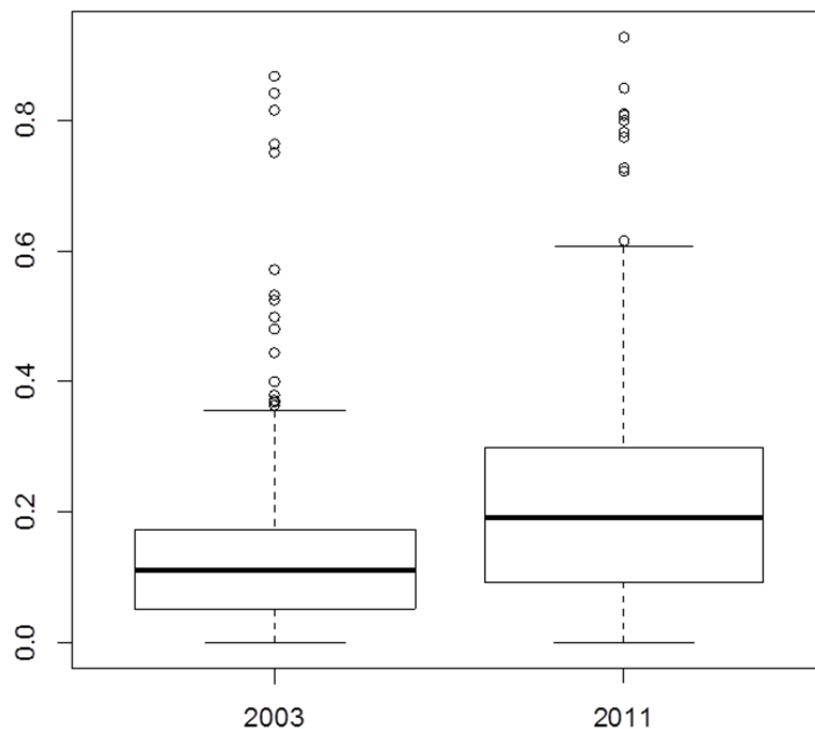
De rupsplots in Grafiek 2 bevestigen deze resultaten. In deze grafiek werd het Vlaams gemiddelde van 2003 respectievelijk 2011 gelijkgesteld met 0.0. De verticale lijn geeft het 95% betrouwbaarheidsniveau aan. Dit wil zeggen dat met 95% zekerheid het gemiddelde van de school zich in het aangegeven bereik bevindt. Deze grafiek toont de verschillen tussen scholen in gemiddeld aantal boeken thuis. Hoewel zowel in 2003 als in 2011 een aantal scholen een significant hoger of lager schoolgemiddelde in aantal boeken thuis (of SES) hebben dan het Vlaamse gemiddelde, wijken de meeste scholen niet significant af van dit Vlaamse gemiddelde, m.a.w. de meeste scholen hebben geen significant andere SES-compositie dan de gemiddelde Vlaamse school. Deze resultaten zijn gelijklopend in beide afnamejaren, hoewel we in 2003 net iets meer scholen zien die zich boven het Vlaams gemiddelde situeren dan in 2011.

Grafiek 2. Verschillen tussen scholen met betrekking tot gemiddeld aantal boeken thuis (in categorieën) in 2003 (boven) en 2011 (onder).



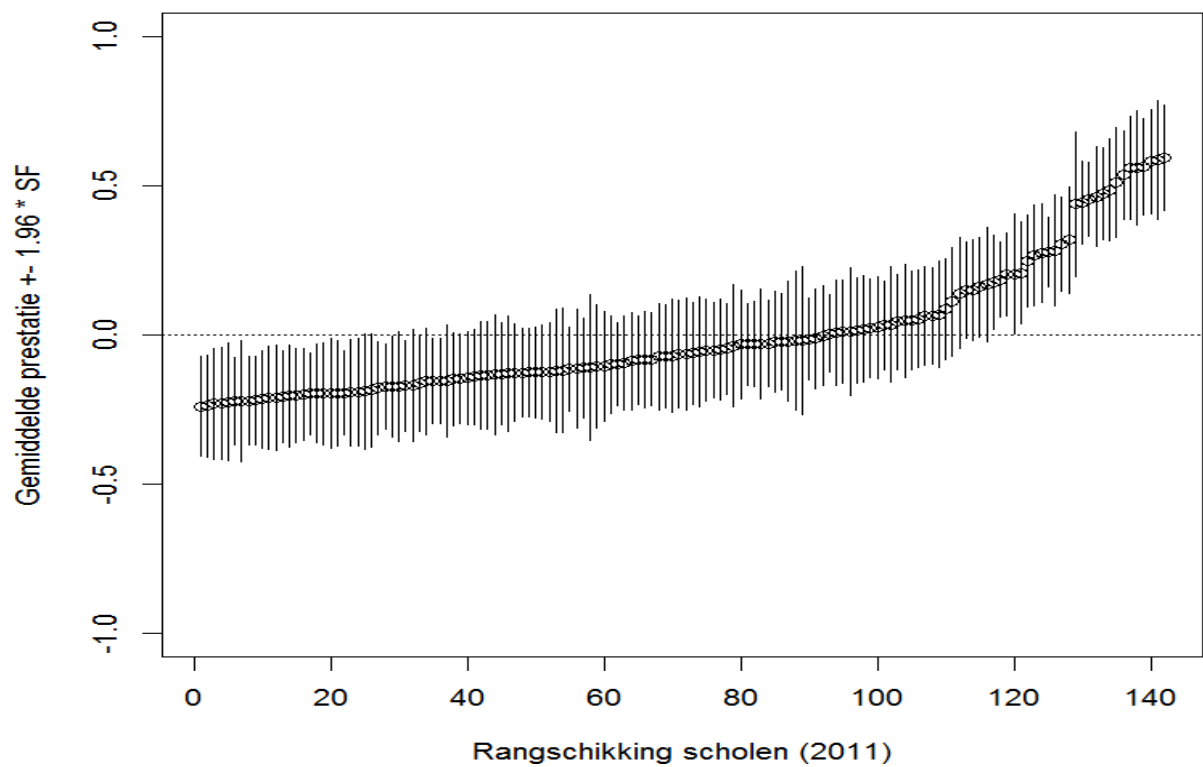
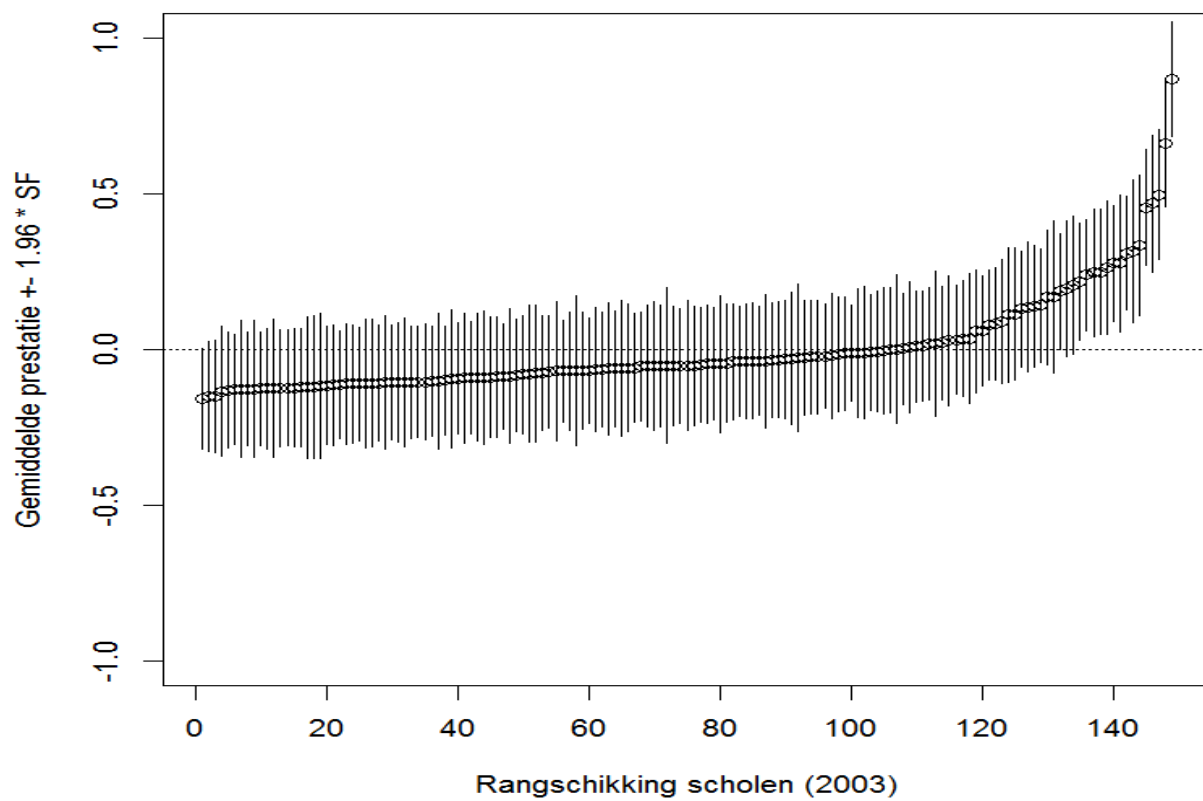
Met betrekking tot het percentage anderstaligen per school zien we grotere verschuivingen (Grafiek 3): niet enkel is het gemiddelde hoger in 2011, ook is er een veel grotere spreiding tussen de scholen. In 2003 hadden ongeveer 75% van de getrokken scholen nog minder dan 20% leerlingen die thuis zelden tot nooit Nederlands praatten. In de steekproef van 2011 heeft 75% van de scholen bijna 30% leerlingen die aangeven zelden tot nooit Nederlands te spreken in hun thuisomgeving. Opvallend zijn ook een aantal scholen met zeer grote uitschieters ($> 70\%$ anderstaligen).

Grafiek 3. Boxplot verdeling % anderstaligen per school, in 2003 en 2011



In Grafiek 4 zien we de verschillen tussen scholen met betrekking tot het percentage anderstaligen op school voor beide afnamejaren. Het percentage anderstaligen is hierbij gecentreerd, waarbij 0.0 het gemiddelde percentage anderstaligen per school in Vlaanderen voorstelt. Hier zien we duidelijke verschillen tussen 2003 en 2011. In 2003 is er geen enkele school die een significant lager percentage anderstaligen heeft dan het Vlaamse gemiddelde, terwijl er wel een aantal scholen zijn dat een significant hoger percentage anderstaligen telt. In 2011 waaiert de rupsplot veel sterker uit: zo zijn er een aantal scholen die een significant hoger percentage anderstaligen op school heeft dan het Vlaamse gemiddelde én scholen met een significant lager percentage anderstaligen op school. Ter verduidelijking: het Vlaams gemiddelde was in 2003 16.54% en in 2011 24.99%; dit gemiddelde wordt in Grafiek 4 telkens gelijkgesteld aan 0.0.

Grafiek 4. Verschillen tussen scholen met betrekking tot percentage anderstaligen op school in 2003 (boven) en 2011 (onder).



2.3. Samenhang tussen achtergrondkenmerken van leerlingen en wiskundeprestaties in 2003 en 2011

In Tabel 3 en Tabel 4 zien we dat er een kloof bestaat tussen zwakke en sterke presteerders. Zo behaalt 5% van de leerlingen een prestatiescore lager dan 451 of 450, respectievelijk in 2003 en 2011, terwijl er ook 5% van de leerlingen een score behaalt die hoger ligt dan 644 of 645, respectievelijk in 2003 en 2011 (Tabel 3). In Tabel 4 zien we verder dat 21% van de leerlingen maximum een laag prestatieniveau bereikt in TIMSS in 2003, ten opzichte van 27% van de leerlingen in 2011, ten opzichte van respectievelijk 26% en 24% van de leerlingen die een hoog tot geavanceerd prestatieniveau behalen. Een kloof tussen zwak en sterk presterende leerlingen is niet noodzakelijk een negatief punt. Dit kan namelijk een indicatie vormen van het feit dat elke leerling de nodige kansen krijgt om zijn/haar maximaal prestatieniveau te bereiken. Wel kan dit mogelijks een probleem vormen wanneer deze kloof tussen sterke en zwakke presteerders samenhangt met de achtergrondkenmerken van de leerling. M.a.w. dat de achtergrondkenmerken van de leerling een samenhang vertonen met het prestatieniveau dat de leerling in kwestie behaalt. Dit betekent dat leerlingen met een andere achtergrond, andere kansen hebben om een hoog prestatieniveau te behalen. Wanneer dit het geval is, spreken we in Vlaanderen over ongelijke kansen. Wanneer leerprestaties samenhangen met de socio-economische status spreken we over sociale ongelijkheid; wanneer deze leerprestaties samenhangen met de thuistaal die de leerling spreekt, spreken we over etnische ongelijkheid (in dit rapport benaderd op basis van de thuistaal van de leerling). In deze paragraaf gaan we op beschrijvende wijze na in welke mate deze achtergrondkenmerken een samenhang vertonen met wiskundeprestaties van leerlingen.

Tabel 3. Wiskundeprestaties, opgedeeld naar percentielen.

	2003					2011				
	5 ^{de}	25 ^{ste}	50 ^{ste}	75 ^{ste}	95 ^{ste}	5 ^{de}	25 ^{ste}	50 ^{ste}	75 ^{ste}	95 ^{ste}
Wiskunde	451.09	510.90	551.54	592.12	643.58	449.97	508.88	549.96	589.84	645.36

Tabel 4. Wiskundeprestaties, opgedeeld naar prestatieniveau in TIMSS.

	2003					2011				
	< Laag	Laag	Middelmatig	Hoog	Geavanceerd	< Laag	Laag	Middelmatig	Hoog	Geavanceerd
% leerlingen	2	19	51	24	2	4	23	49	22	2

Tabel 5 toont de correlaties tussen een aantal achtergrondkenmerken van leerlingen en hun wiskundeprestaties in 2003. Hierbij is te zien dat wiskundeprestaties niet significant samenhangen met het geslacht van de leerling, maar wel met de geboortemaand, in de zin dat leerlingen die later op het jaar geboren zijn, lagere prestaties behalen. Ook vertonen de wiskundeprestaties een samenhang met het al dan niet normaalvorderend zijn, waarbij leerlingen die versneld of normaalvorderend zijn hogere prestaties behalen dan leerlingen die een schoolse achterstand hebben opgelopen. Vervolgens zien we ook dat leerlingenprestaties samenhangen met de thuistaal van de leerlingen (als indicator van etniciteit) en het aantal boeken thuis (als indicator van SES). Leerlingen die thuis altijd Nederlands spreken behalen hogere wiskundeprestaties dan leerlingen die thuis soms Nederlands spreken, en deze leerlingen behalen op hun beurt hogere prestaties dan leerlingen die thuis nooit Nederlands spreken. Met betrekking tot het aantal boeken thuis, geldt dat leerlingen met een hoger aantal boeken thuis, een hogere prestatie hebben.

In 2011 (Tabel 6) zien we soortgelijke resultaten; ook hier bestaat er een significante samenhang tussen wiskundeprestaties enerzijds en geboortemaand, normaalvorderend zijn, thuistaal en aantal boeken anderzijds. Wel zien we dat ook geslacht in dit afnamejaar een significante samenhang vertoont met de wiskundeprestaties, in de zin dat jongens significant hogere wiskundeprestaties behalen dan meisjes.

Tabel 5. Correlatie tussen geslacht, geboortemaand, al dan niet normaalvorderend, thuistaal, aantal boeken thuis en wiskundeprestaties, in 2003.

	Geslacht	Geboortemaand	Normaalvorderend	Thuistaal	Aantal boeken	Wiskundeprestaties
Geslacht ^a	1.00					
Geboortemaand ^b	-.02 [*]	1.00				
Normaalvorderend ^c	-.04 ^{**}	-.14 ^{***}	1.00			
Thuistaal ^d	.09 ^{***}	.03 ^{***}	-.14 ^{***}	1.00		
Aantal boeken ^e	-.01	-.03	.20 ^{***}	-.08 ^{***}	1.00	
Wiskundeprestaties	.02	-.07 ^{***}	.35 ^{***}	-.20 ^{***}	.22 ^{***}	1.00

* p < .05; ** p < .01; *** p < .001

Tabel 6. Correlatie tussen geslacht, geboortemaand, al dan niet normaalvorderend, thuistaal, aantal boeken thuis en wiskundeprestaties, in 2011.

	Geslacht	Geboortemaand	Normaalvorderend	Thuistaal	Aantal boeken	Wiskunde
Geslacht ^a	1.00					
Geboortemaand ^b	-.02	1.00				
Normaalvorderend ^c	.01	-.14 ^{***}	1.00			
Thuistaal ^d	.00	.02	-.17 ^{***}	1.00		
Aantal boeken ^e	-.01	-.04 [*]	.18 ^{***}	-.11 ^{***}	1.00	
Wiskunde	.07 ^{***}	-.09 ^{***}	.35 ^{***}	-.19 ^{***}	.20 ^{***}	1.00

* p < .05; ** p < .01; *** p < .001

^a 0 = meisje, 1 = jongen

^b 0 = januari; 1 = februari; 2 = maart; 3 = april; 4 = mei; 5 = juni; 6 = juli; 7 = Augustus; 8 = september; 9 = oktober; 10 = november; 11 = december

^c -2 = 2 jaar of meer vertraging; -1 = 1 jaar vertraging; 0 = normaalvorderend; 1 = versneld

^d 1 = altijd (of bijna) Nederlands; 2 = soms Nederlands; 3 = nooit Nederlands

^e 0 = 0-10 boeken; 1 = 11-25 boeken; 2 = 26-100 boeken; 3 = 101-200 boeken; 4 = meer dan 200 boeken

De correlatietabellen geven een eerste indicatie dat er sprake is van ongelijkheden in het Vlaamse onderwijs. Om de ongelijkheid op basis van SES en thuistaal meer in detail te bekijken, geven we in Tabel 7 het gemiddelde van de leerlingkenmerken thuistaal en aantal boeken thuis weer, opgedeeld per prestatieniveau in TIMSS.

Tabel 7. Gemiddelde thuistaal en gemiddeld aantal boeken thuis (in categorieën) per prestatieniveau in TIMSS.

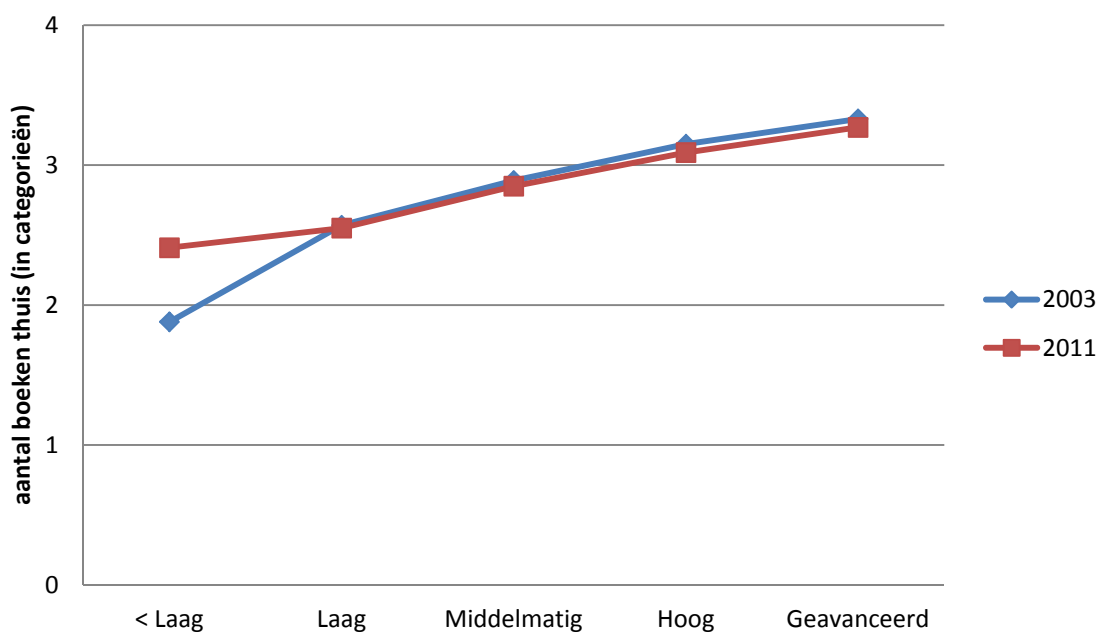
	2003					2011				
	< Laag	Laag	Middelmatig	Hoog	Geavanceerd	< Laag	Laag	Middelmatig	Hoog	Geavanceerd
Thuistaal ^a	1.71	1.37	1.23	1.15	1.07	1.64	1.46	1.33	1.21	1.14
Aantal boeken ^b	1.88	2.57	2.89	3.15	3.33	2.41	2.55	2.85	3.09	3.27

^a 1 = altijd (of bijna) Nederlands; 2 = soms Nederlands; 3 = nooit Nederlands

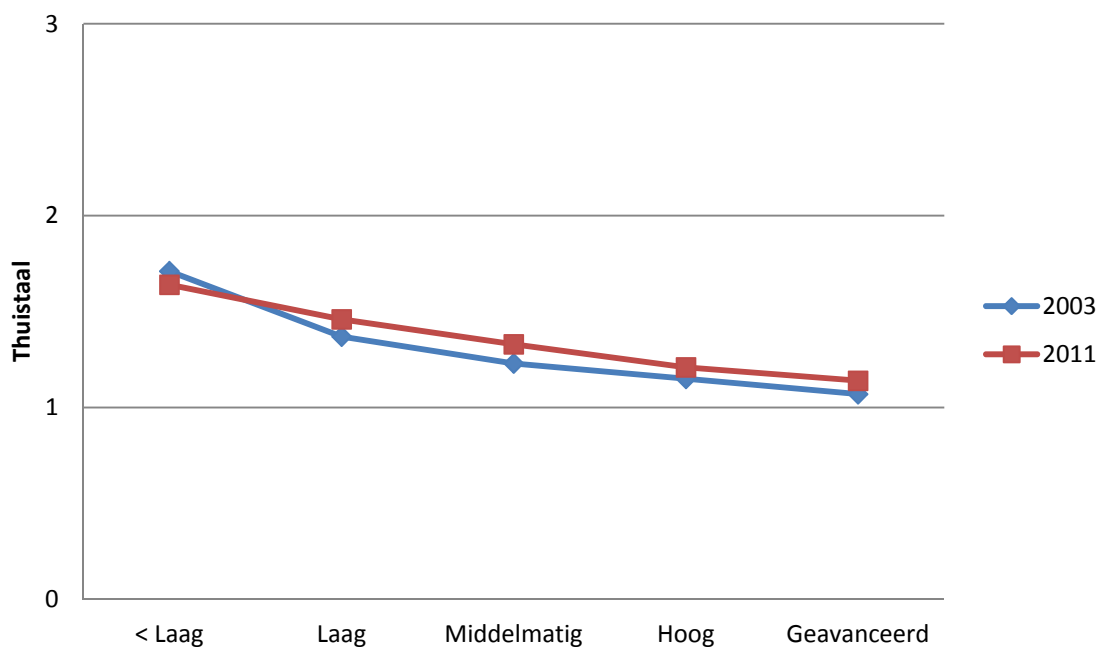
^b 1 = 0-10 boeken; 2 = 11-25 boeken; 3 = 26-100 boeken; 4 = 101-200 boeken; 5 = meer dan 200 boeken

Als we Tabel 7 grafisch voorstellen in Grafiek 5 en Grafiek 6, zien we de samenhang tussen de prestatieniveau in TIMSS en het gemiddelde aantal boeken thuis en de thuistaal. Grafiek 5 laat zien dat er een lichte stijging is van het aantal boeken thuis naarmate het prestatieniveau van de leerling stijgt. In 2003 zien we dat voornamelijk de leerlingen die het laagste prestatieniveau niet behalen, een erg laag aantal boeken thuis (11-25 boeken) hebben. In 2011 is dit niet meer het geval. Ook wat betreft thuistaal zien we dat dit samenhangt met de wiskundeprestaties, in die zin dat er een trend bestaat dat de leerlingen die een hoger prestatieniveau behalen in TIMSS, ook vaker thuis altijd of soms Nederlands spreken. Leerlingen met een lager prestatieniveau, spreken gemiddeld vaker slechts soms of nooit Nederlands thuis. Deze gegevens geven aan dat er wel degelijk indicaties zijn dat de wiskundeprestaties in het Vlaamse onderwijs althans gedeeltelijk verklaard kunnen worden door middel van achtergrondkenmerken van leerlingen, met name met hun thuistaal (als indicator van etniciteit) en het aantal boeken thuis (SES). Globaal beschouwd is het verschil op die punten tussen 2011 en 2003 klein. In Deel 3 van dit rapport gaan we na in welke mate er sprake is van significante samenhang tussen deze achtergrondkenmerken en wiskundeprestaties, ook wanneer gecontroleerd wordt voor een aantal andere achtergrondkenmerken van leerlingen. Eveneens bekijken we in welke mate deze mogelijke ongelijkheden zijn gewijzigd tussen 2003 en 2011.

Grafiek 5. Gemiddeld aantal boeken thuis (in categorieën) opgedeeld naar prestatieniveau in TIMSS, in 2003 en 2011



Grafiek 6. Gemiddelde thuistaal opgedeeld naar prestatieniveau in TIMSS, in 2003 en 2011



3. Ongelijkheid op basis van SES en thuistaal in wiskundeprestaties

In dit deel voeren we statistische toetsen uit om na te gaan in welke mate de SES en de thuistaal van leerlingen en de schoolcompositie in termen van SES en het percentage anderstaligen per school significant samenhangen met de wiskundeprestaties in Vlaanderen. M.a.w. we gaan na in welke mate er sprake is van sociale ongelijkheid en ongelijkheid op basis van thuistaal in het Vlaamse wiskundeonderwijs. In een eerste paragraaf gaan we na in welke mate er globaal gezien sprake is van ongelijkheid op basis van SES en thuistaal in het Vlaamse wiskundeonderwijs en doen we een uitspraak over de mate waarin deze ongelijkheden zijn gewijzigd in 2011 ten opzichte van 2003. Vervolgens gaan we in de tweede paragraaf na of er verschillen zijn tussen scholen in de mate waarin SES en thuistaal samenhangen met wiskundeprestaties en of die verschillen tussen scholen toegenomen of afgenomen zijn.

De analyses maken gebruik van de vijf 'plausibele waarden' die ter beschikking werden gesteld door TIMSS als prestatie-indicatoren (zie Bijlage 4).

3.1. Globale ongelijkheid op basis van SES en thuistaal in het Vlaamse wiskundeonderwijs

In deze paragraaf wordt eerst bekeken in welke mate er sprake is van ongelijkheid op basis van SES en thuistaal in het Vlaamse wiskundeonderwijs, respectievelijk in 2003 en in 2011. Vervolgens gaan we na in welke mate deze (on)gelijkheid gewijzigd is in 2011 ten opzichte van 2003.

3.1.1. Invloed van achtergrondkenmerken en schoolcompositie op wiskundeprestaties in 2003 en 2011

De onderzoeksvragen die in deze paragraaf centraal staan, luiden:

- In welke mate is er een samenhang tussen het aantal boeken thuis van de leerling en diens wiskundeprestaties, respectievelijk in 2003 en 2011, nadat gecontroleerd wordt voor geslacht, geboortemaand, al dan niet normaalvorderend zijn en thuistaal?
- In welke mate is er een samenhang tussen de thuistaal van de leerling en diens wiskundeprestaties, respectievelijk in 2003 en 2011, nadat gecontroleerd wordt voor geslacht, geboortemaand, al dan niet normaalvorderend zijn en aantal boeken thuis?
- In welke mate is er een samenhang tussen het schoolgemiddelde in aantal boeken thuis en de wiskundeprestaties van leerlingen, respectievelijk in 2003 en 2011, nadat gecontroleerd wordt voor de leerlingkenmerken geslacht, geboortemaand, al dan niet normaalvorderend zijn, aantal boeken thuis en thuistaal en ook voor de schoolcompositie inzake thuistaal?
- In welke mate is er een samenhang tussen het percentage anderstaligen op school en de wiskundeprestaties van leerlingen, respectievelijk in 2003 en 2011, nadat gecontroleerd wordt voor de leerlingkenmerken geslacht, geboortemaand, al dan niet normaalvorderend zijn, aantal boeken thuis en thuistaal en ook voor de schoolcompositie inzake SES?

Om een antwoord te bieden op deze vragen, werden multilevel analyses uitgevoerd, afzonderlijk op de data van TIMSS 2003 en TIMSS 2011. Doordat bij TIMSS volledige klassen in het onderzoek opgenomen worden en er vaak meerdere klassen per school deelnemen, kunnen er drie niveaus onderscheiden worden, met name een school-, klas- en leerlingniveau. De analyses houden hiermee rekening door middel van drielevel multilevel-analyses, waarbij leerlingen genest zijn binnen klassen en klassen binnen scholen. De resultaten van deze multilevel-analyses staan samengevat in Tabel 9 (zie p. 26).

Allereerst wordt een nulmodel geschat. Dit geeft een zicht op het percentage van de variantie in leerlingresultaten dat gesitueerd is op elk van de drie verschillende niveaus. Deze resultaten zijn opgenomen in Tabel 8. Uit de intra-class correlatie (ICC) blijkt dat ongeveer 13% van de variantie in wiskundeprestaties tussen leerlingen zich op schoolniveau bevindt, ongeveer 4% op klasniveau en de overige 82% zijn verschillen tussen individuele leerlingen binnen klassen. Deze verdeling is zeer gelijkaardig in beide jaren. Concreet betekent dit dat de keuze van de school de score van de leerling in 95% van de gevallen met maximaal 41 à 49 punten (respectievelijk in 2003 en 2011) op de prestatieschaal van TIMSS doet afwijken van het Vlaamse gemiddelde (551.36). De variantie op

leerlingenniveau doet deze score schommelen met 105 à 106 punten. Met andere woorden, de score van leerlingen kan in 95% van de gevallen schommelen tussen 445.36 en 657.36.

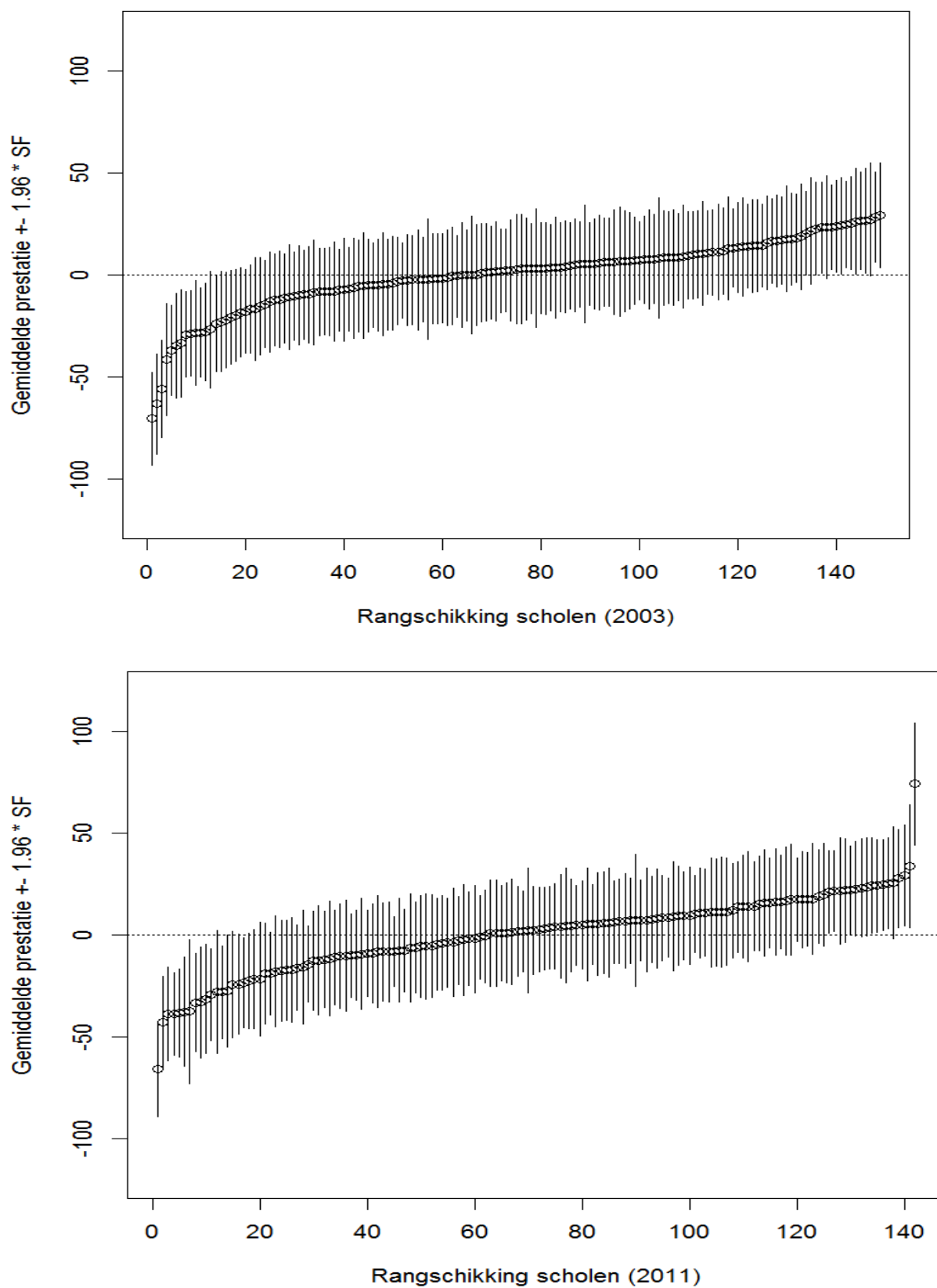
Tabel 8. Proportie van de variantie op school-, klas- en leerlingniveau in 2003 en 2011 (nulmodel).

	2003				2011			
	N	Variantie	S.E.	ICC (%)	N	Variantie	S.E.	ICC (%)
Schoolniveau	149	436.41	97.50	12.56	142	501.66	97.96	14.00
Klasniveau	258	129.17	58.65	3.72	268	156.03	54.81	4.35
Leerlingniveau	4587	2908.00	75.18	83.72	4767	2926.79	80.70	81.65

Grafiek 7 geeft het gemiddelde prestatieniveau van elke school weer met een bolletje. De verticale lijn geeft het 95% betrouwbaarheidsniveau aan. Dit wil zeggen dat met 95% zekerheid het gemiddelde van de school zich in het aangegeven bereik bevindt. De horizontale stippellijn door 0 representeert de gemiddelde wiskundescore van de Vlaamse scholen. Met name werd de gemiddelde Vlaamse score omgezet naar nul. Scholen waarvan het betrouwbaarheidsinterval deze horizontale stippellijn niet kruist, hebben met andere woorden een gemiddeld prestatieniveau voor wiskunde dat significant lager of hoger ligt dan het Vlaamse gemiddelde. Als we kijken naar Grafiek 7 zien we dat een aantal scholen significant boven het gemiddelde (scholen boven de stippellijn) en een aantal scholen significant onder het Vlaamse gemiddelde (scholen onder de stippellijn) prestatieniveau voor wiskunde scoren.

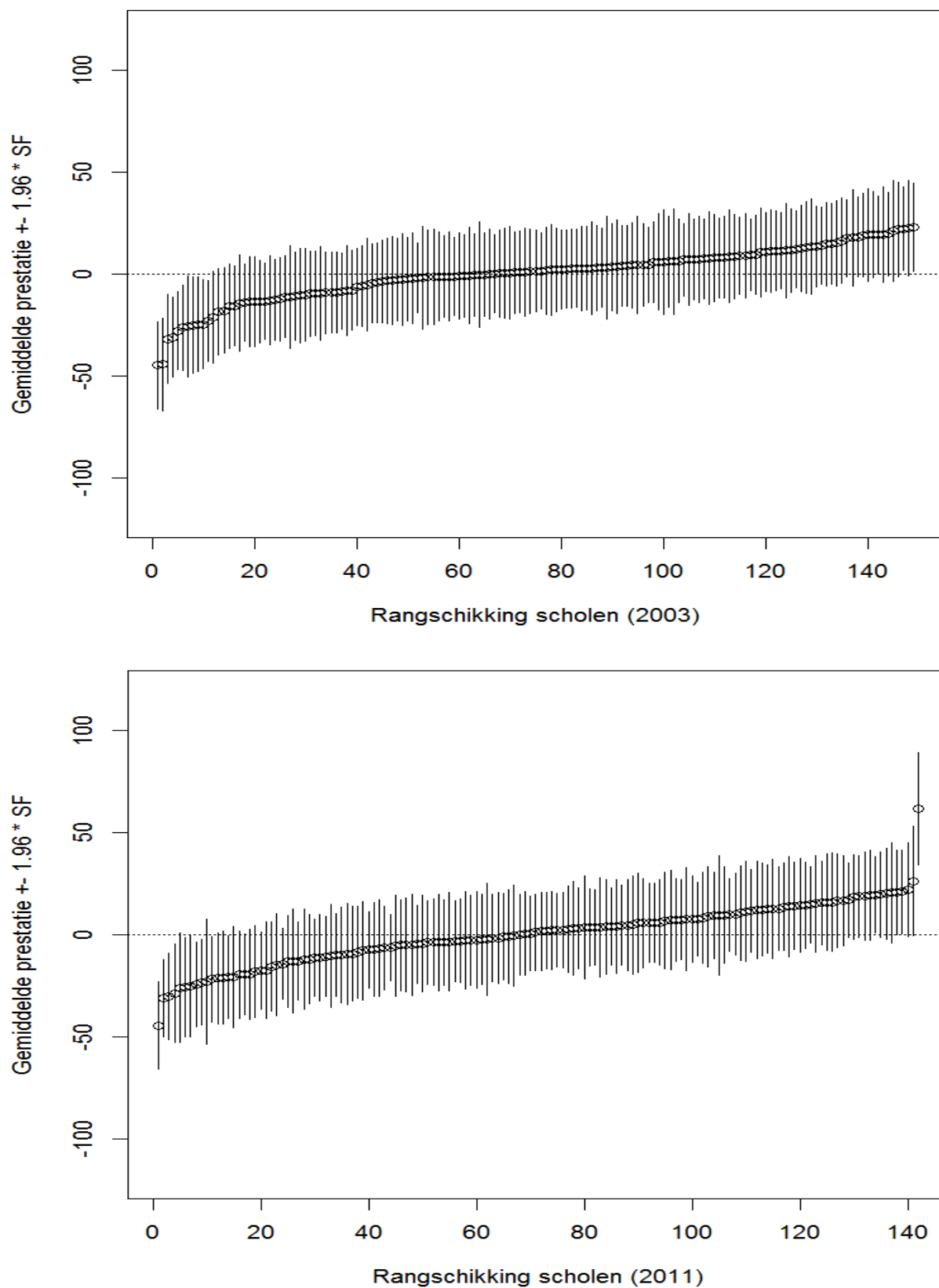
Opvallend is de uitschieter in 2011. Er is met name in 2011 één school die een opvallend hoge wiskundeprestatie behaalt, significant hoger dan het Vlaamse gemiddelde en significant hoger dan het gemiddelde van een aanzienlijk aantal scholen. Hoewel dit ervoor zorgt dat de gemiddelde wiskundeprestatie van de gehele steekproef wat hoger ligt dan wanneer deze school uit de analyses verwijderd wordt, hebben we er toch voor gekozen om deze school in de analyses mee op te nemen. Enerzijds omdat er geen (andere) aanwijzingen zijn dat er zaken zouden zijn misgelopen in de dataverzameling die de scores kunnen beïnvloeden hebben. Anderzijds omdat we de resultaten van deze analyses hebben vergeleken met de analyses in Belfi et al. (2012), waarin de outlier-school wel uit de analyses werd verwijderd, en we zien geen grote afwijkingen tussen beide reeksen resultaten.

Grafiek 7. Rangschikking scholen uit TIMSS 2003 (boven) en TIMSS 2011 (onder) op basis van de gemiddelde wiskundeprestatie, zonder controle voor leerlingkenmerken.



Scholen kunnen echter niet zonder meer verantwoordelijk gesteld worden voor deze verschillen in prestatieniveau. Er zijn namelijk factoren waarop scholen geen invloed hebben, maar die wel een invloed op prestaties hebben. Te denken valt onder meer aan het geslacht van de leerlingen. Scholen hebben een verschillend leerlingenpubliek en dit leerlingenpubliek zal mede het gemiddelde prestatieniveau van een school bepalen. Om een meer eerlijke vergelijking van scholen te bekomen, is het dan ook belangrijk om te controleren voor een aantal kenmerken van dit leerlingenpubliek (Van Damme et al., 2004). Daarnaast zorgt de opname van een aantal leerlingenkenmerken ervoor dat we een zuiverdere schatting krijgen van de invloed van de factoren waarin we geïnteresseerd zijn, met name SES, thuistaal en schoolcompositie in termen van SES en thuistaal. Daarom wordt in Model 1 (zie Tabel 9) een aantal kenmerken van leerlingen toegevoegd, met name geslacht, geboortemaand en al dan niet normaalvorderend zijn. Uit de literatuur komt duidelijk naar voren dat deze kenmerken een invloed hebben op de prestaties van leerlingen (zie o.a. Belfi et al., 2012). Dit komt ook naar voren in Tabel 5 en Tabel 6, waarin we zien dat de voorgenoemde drie kenmerken een significante correlatie vertonen met de wiskundeprestaties (uitgezonderd voor geslacht in afnamejaar 2003). In Grafiek 8 zien we dat, door toevoeging van het geslacht, geboortemaand en al dan niet normaalvorderend zijn, de verschillen tussen scholen afnemen (in vergelijking met de resultaten in Grafiek 7), zowel in 2003 als in 2011. Dit betekent dat een deel van de verschillen tussen scholen in wiskundeprestaties verklaard kan worden door de verschillen die er bestaan tussen de leerlingengroepen per school naar geslacht, geboortemaand en al dan niet normaalvorderend zijn.

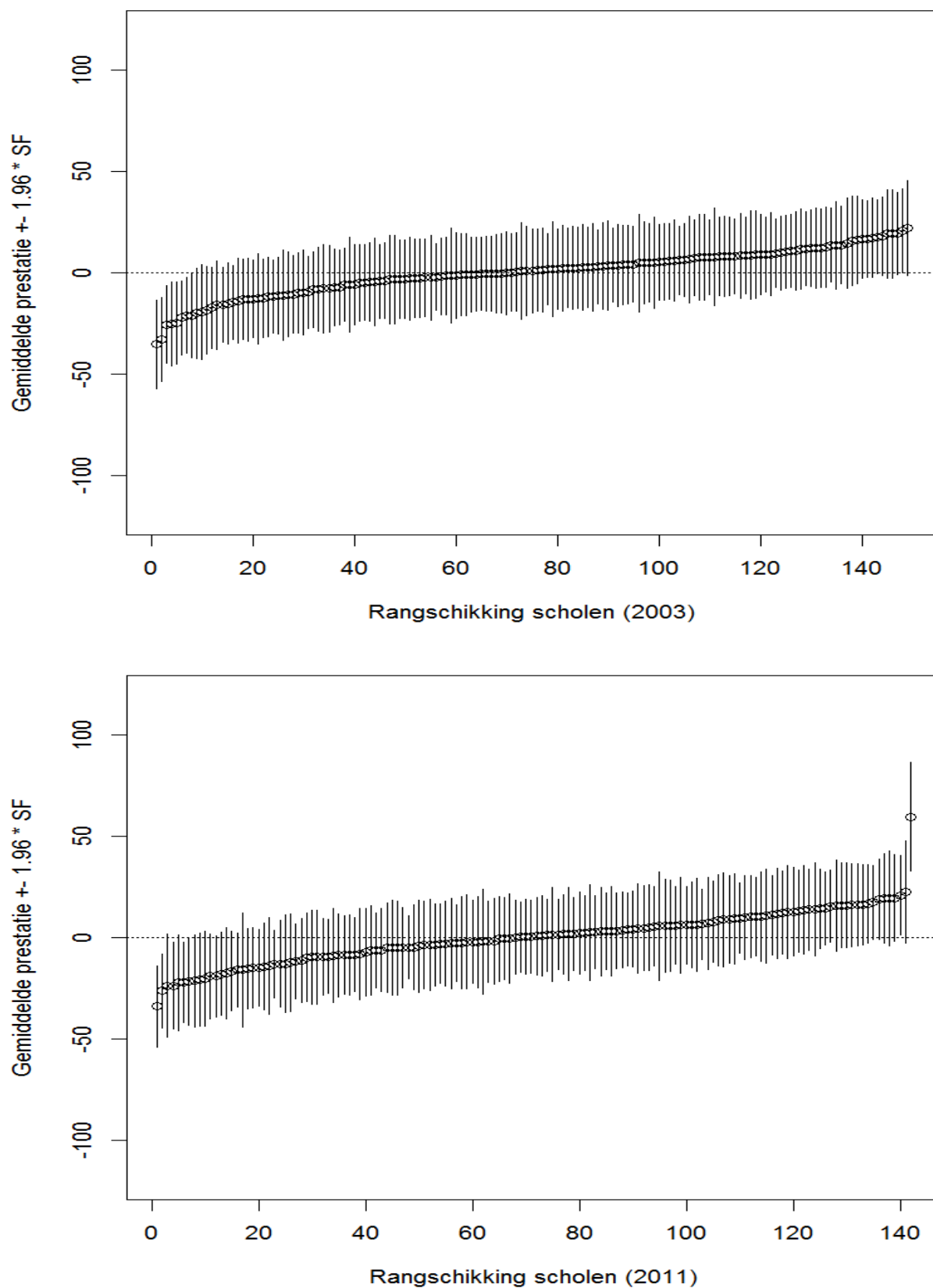
Grafiek 8. Rangschikking scholen uit TIMSS 2003 (boven) en TIMSS 2011 (onder) op basis van de gemiddelde wiskundeprestatie, gecontroleerd voor de leerlingkenmerken geslacht, geboortemaand en al dan niet normaalvorderend zijn.



Model 1 in Tabel 9 toont de resultaten van de analyses waarbij gecontroleerd wordt voor geslacht, geboortemaand en al dan niet normaalvorderend, respectievelijk voor 2003 en 2011. De resultaten vertonen gelijkenissen in beide afnamejaren. Zowel in 2003 als in 2011 is er een significant effect van geslacht; jongens behalen met name in beide afnamejaren een significant hogere wiskundeprestatie dan meisjes (na controle voor geboortemaand en al dan niet normaalvorderend zijn). Ten tweede is er in beide afnamejaren een significant negatief effect van zittenblijven, in die zin dat leerlingen die normaalvorderend zijn hogere wiskundeprestaties behalen dan leerlingen die minstens eenmaal zijn blijven zitten, en dat versnelde leerlingen hogere wiskundeprestaties behalen dan leerlingen die normaalvorderend zijn (steeds na controle voor geslacht en geboortemaand). De resultaten met betrekking tot geboortemaand zijn verschillend in 2003 en 2011. Waar in 2011 een significant effect bestaat van geboortemaand, in de zin dat leerlingen die later op het jaar geboren zijn, lagere wiskundeprestaties behalen, is deze samenhang tussen geboortemaand en wiskundeprestaties niet significant in 2003 (steeds na controle voor geslacht en al dan niet normaalvorderend zijn). Deze kenmerken kunnen in beide afnamejaren ca. 13% van de verschillen in leerlingprestaties verklaren.

In een volgende model (Model 2, Tabel 9) wordt gekeken in welke mate het aantal boeken thuis en de thuistaal een samenhang vertonen met de wiskundeprestaties van leerlingen, rekening houdend met het geslacht, de geboortemaand en het al dan niet normaalvorderend zijn van de leerling. Ook hier zijn de resultaten erg gelijklopend in beide afnamejaren. Allereerst zien we ook hier dat de opname van deze leerlingkenmerken ervoor zorgt dat de verschillen tussen scholen in wiskundeprestaties nog kleiner worden (zie Grafiek 9, in vergelijking met Grafiek 8).

Grafiek 9. Rangschikking scholen uit TIMSS 2003 (boven) en TIMSS 2011 (onder) op basis van de gemiddelde wiskundeprestatie, gecontroleerd voor leerlingenkenmerken geslacht, geboortemaand, normaalvorderend, aantal boeken thuis en thuistaal.

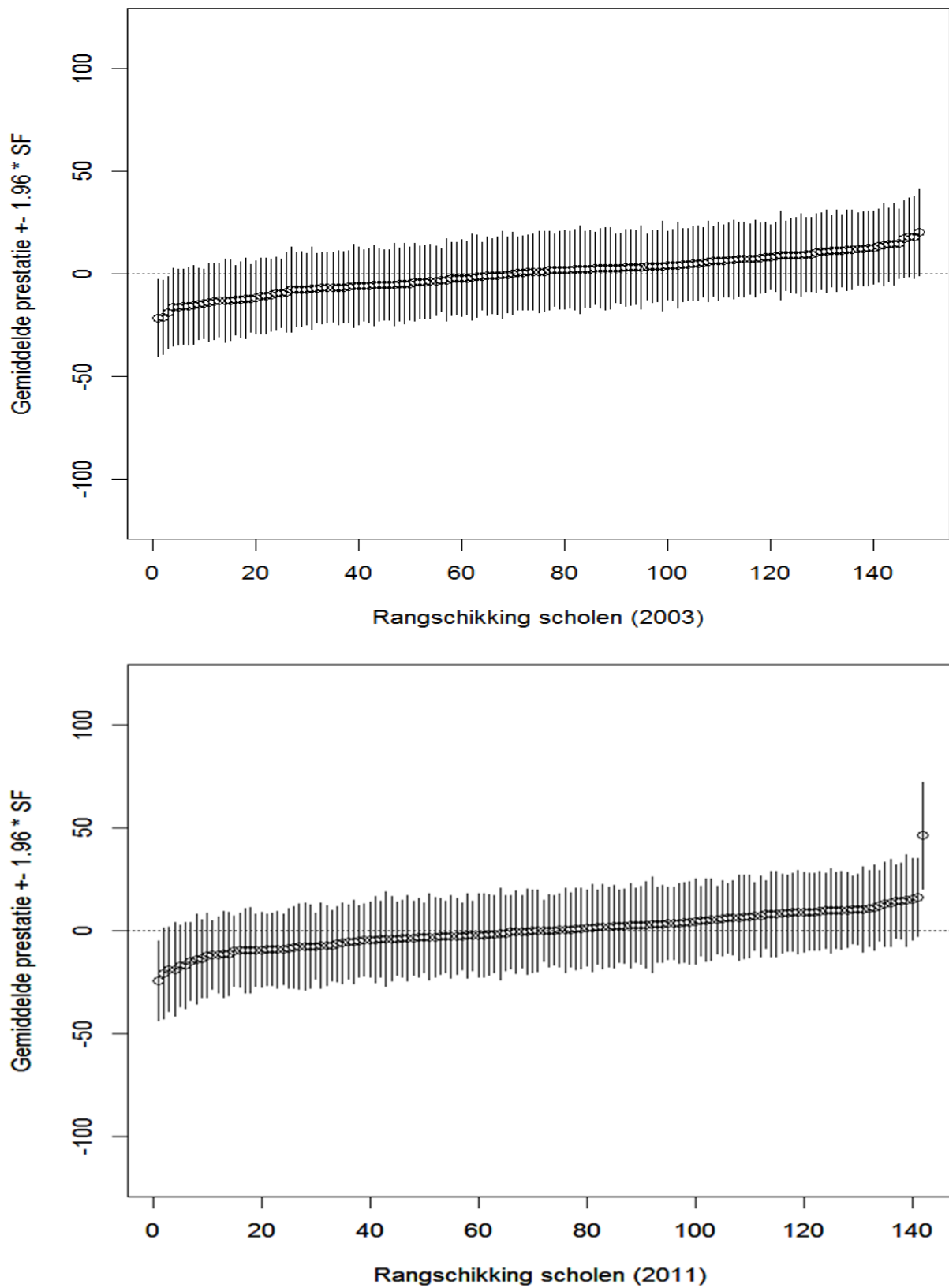


Daarnaast is in Model 2 in Tabel 9 te zien dat zowel het aantal boeken thuis als de thuistaal van de leerling een significante samenhang vertonen met diens wiskundeprestaties, zowel in 2003 als in 2011 en dit na controle voor geslacht, geboortemaand en al dan niet normaalvorderend zijn. Met name behalen leerlingen met minder boeken thuis, significant lagere wiskundeprestaties, na controle voor geslacht, geboortemaand, al dan niet normaalvorderend zijn en thuistaal. Leerlingen die thuis nooit Nederlands spreken behalen significant lagere wiskundeprestaties dan leerlingen die thuis soms Nederlands spreken, en deze leerlingen behalen op hun beurt lagere wiskundeprestaties dan leerlingen die thuis altijd Nederlands spreken. Beide kenmerken zorgen ervoor dat er bijkomend ca. 3,50% van de variantie in leerlingprestaties kan verklaard worden. We willen er hier op wijzen dat de geschatte coëfficiënten van 2003 en 2011 verschillen vertonen, voornamelijk bij thuistaal. Dit lijkt erop te wijzen dat de thuistaal in 2003 een grotere negatieve impact heeft dan in 2011 (in 2003 behaalt een leerling die soms Nederlands spreekt gemiddeld een score die 11.48 punten lager ligt dan de score van een leerling die thuis altijd Nederlands spreekt; in 2011 is dit nog 9.72 punten – op de prestatieschaal van TIMSS, met internationaal gemiddelde van 500 en standaardafwijking van 100). Zoals reeds gezegd, is het in deze analyses echter niet mogelijk om uitspraken te doen over de significantie van dit verschil tussen beide afnamemomenten. Dit wordt onderzocht in paragraaf 3.1.2.

In een laatste model (Model 3, Tabel 9) bekijken we in welke mate de schoolcompositie inzake SES en het percentage anderstaligen op school een samenhang vertonen met de wiskundeprestaties van leerlingen. Opnieuw zijn de resultaten erg gelijklopend in beide afnamejaren. De schoolcompositie inzake SES en thuistaal zorgt ervoor dat bijkomend ca. 2% van de verschillen in leerlingprestaties verklaard kunnen worden. Dit betekent dat alle opgenomen kenmerken ca. 18,50% van de variantie in leerlingprestaties verklaard wordt. De schoolcompositie inzake SES vertoont in beide gevallen geen significante samenhang met de wiskundeprestaties (na controle voor geslacht, geboortemaand, al dan niet normaalvorderend zijn, aantal boeken thuis, thuistaal en percentage anderstaligen). Het percentage anderstaligen op school vertoont wel een significante negatieve samenhang met de wiskundeprestaties van leerlingen, in die zin dat een hoger percentage anderstaligen samenhangt met lagere wiskundeprestaties van leerlingen (steeds gecontroleerd voor voornoemde variabelen). Een vergelijking van de coëfficiënten wijst in de richting dat dit effect kleiner is geworden in 2011 in vergelijking met 2003, maar door middel van deze analyses kunnen geen uitspraken gedaan worden over de significantie van dit verschil. De opname van deze twee schoolkenmerken zorgt er bijkomend voor dat de verschillen tussen scholen verkleinen. Meer bepaald toont Grafiek 10 dat er in 2003 nog slechts drie scholen zijn die significant lagere resultaten behalen dan het Vlaamse prestatiegemiddelde voor wiskunde als rekening gehouden wordt met de vermelde variabelen; in

2011 behaalt één school significant hogere resultaten en één school significant lagere resultaten voor wiskunde dan het Vlaamse gemiddelde.

Grafiek 10. Rangschikking scholen uit TIMSS 2003 (boven) en TIMSS 2011 (onder) op basis van de gemiddelde wiskundeprestatie, gecontroleerd voor leerlingenkenmerken (geslacht, geboortemaand, normaalvorderend, aantal boeken thuis en thuistaal) en schoolcompositie inzake SES en thuistaal.



Tabel 9. Invloed aantal boeken thuis, thuistaal en schoolcompositie op wiskundeprestaties in 2003.

	2003								2011							
	Model 0		Model 1		Model 2		Model 3		Model 0		Model 1		Model 2		Model 3	
	b	S.E.	b	S.E.	b	S.E.	b	S.E.	b	S.E.	b	S.E.	b	S.E.	b	S.E.
Intercept	551.36***	1.47	557.71***	2.23	543.40***	2.96	534.59***	8.77	550.59***	1.69	557.00***	3.01	544.64***	3.61	543.21***	8.39
<i>Leerlingniveau</i>																
Geslacht ^a			4.96**	1.80	5.52**	1.74	5.70**	1.74			7.81**	2.16	8.11***	2.16	8.06***	2.15
Geboortemaand ^b			-.28	.26	-.24	.26	-.25	.26			-.76*	.37	-.71	.36	-.73	.37
Normaaltvorderend ^c			42.25***	2.26	37.76***	2.30	37.28***	2.27			38.61***	2.17	35.18***	2.15	34.77***	2.19
Aantal boeken thuis ^d					7.72***	.99	7.45***	1.03					7.23***	1.03	7.04***	1.04
Thuistaal ^e					-11.48***	2.23	-9.68***	2.19					-9.72***	2.17	-7.99***	2.19
<i>Schoolniveau</i>																
Compositie SES ^f							7.94	4.20							5.32	4.27
Compositie thuistaal ^g							-43.94***	6.56							-36.31***	6.99
<i>Variantie (random gedeelte)</i>																
School (intercept)	436.41	97.50	283.46	64.42	227.06	52.64	166.43	40.04	501.66	97.96	336.35	78.11	275.36	76.71	176.87	74.12
Klas (intercept)	129.17	58.65	129.20	47.41	126.71	45.59	120.28	43.21	156.03	54.81	149.21	47.25	144.11	49.52	168.54	57.03
Leerling	2907.99	75.18	2624.13	71.48	2545.68	72.27	2546.01	72.21	2926.79	80.70	2639.94	77.91	2576.22	77.07	2573.87	77.07
Totale proportie verklaarde variantie			12.57		16.53		18.45				12.80		16.43		18.56	
Log Likelihood	-24963.57		-24715.20***		-24637.3***		-24623.2***		-25966.20		-25708.00***		-25641.40***		-25629.00***	

* p < .05; ** p < .01; *** p < .001

^a 0 = meisje, 1 = jongen^b 0 = januari; 1 = februari; 2 = maart; 3 = april; 4 = mei; 5 = juni; 6 = juli; 7 = Augustus; 8 = september; 9 = oktober; 10 = november; 11 = december^c -2 = 2 jaar of meer vertraging; -1 = 1 jaar vertraging; 0 = normaalvorderend; 1 = versneld^d 0 = 0-10 boeken; 1 = 11-25 boeken; 2 = 26-100 boeken; 3 = 101-200 boeken; 4 = meer dan 200 boeken^e 1 = altijd (of bijna) Nederlands; 2 = soms Nederlands; 3 = nooit Nederlands^f Geaggregeerde variabele van het aantal boeken thuis. Per school werd het gemiddelde berekend van de variabele aantal boeken thuis.^g Percentage anderstaligen per school. Per school werd het percentage leerlingen berekend dat aangaf thuis nooit of soms Nederlands te spreken.

3.1.2. Veranderingen tussen 2003 en 2011 inzake invloed van achtergrondkenmerken van leerlingen en schoolcompositie op wiskundeprestaties

In de vorige paragraaf kregen we een zicht op de samenhang tussen een aantal leerling- en schoolkenmerken en de prestaties van leerlingen in wiskunde, afzonderlijk voor 2003 en 2011. Hoewel dit een indicatie biedt van mogelijke verschillen tussen beide afnamejaren, kan op basis van deze analyses niet besloten worden in welke mate de mogelijke verschillen tussen beide afnamejaren significant zijn. Daarom worden in deze paragraaf de analyses verder uitgebreid en bieden we een antwoord op de onderzoeksvragen:

- In welke mate is de samenhang tussen het aantal boeken thuis van de leerling en diens wiskundeprestaties, gewijzigd tussen 2003 en 2011, nadat gecontroleerd wordt voor geslacht, geboortemaand, al dan niet normaalvorderend zijn en thuistaal?
- In welke mate is de samenhang tussen de thuistaal van de leerling en diens wiskundeprestaties, gewijzigd tussen 2003 en 2011, nadat gecontroleerd wordt voor geslacht, geboortemaand, al dan niet normaalvorderend zijn en aantal boeken thuis?
- In welke mate is de samenhang tussen het schoolgemiddelde in aantal boeken thuis en de wiskundeprestaties van leerlingen gewijzigd tussen 2003 en 2011, nadat gecontroleerd wordt voor de leerlingkenmerken geslacht, geboortemaand, al dan niet normaalvorderend zijn, aantal boeken thuis, thuistaal en schoolcompositie in termen van thuistaal?
- In welke mate is de samenhang tussen het percentage anderstaligen op school en de wiskundeprestaties van leerlingen gewijzigd tussen 2003 en 2011, nadat gecontroleerd wordt voor de leerlingkenmerken geslacht, geboortemaand, al dan niet normaalvorderend zijn, aantal boeken thuis, thuistaal en schoolcompositie inzake SES?

Om uitspraken te kunnen doen over de mate waarin de onderzochte leerling- en schoolkenmerken een verschillende samenhang vertonen in 2003 en in 2011, nemen we de data van beide afnamejaren tesamen op in één model. Tabel 10 geeft de variantie weer op school-, klas- en leerlingniveau wanneer de data van 2003 en 2011 tesamen bekeken worden. 13.24% van de verschillen in leerlingprestaties kan verklaard worden door de verschillen tussen scholen; 4.05% kan verklaard worden doordat leerlingen in andere klassen zitten binnen een school en 82.71% van de verschillen in leerlingenprestaties is toe te schrijven aan verschillen tussen leerlingen binnen een klas. Hierbij spreken we echter over scholen, klassen en leerlingen die hetzij in 2003, hetzij in 2011 bevraagd werden. Omdat we voornamelijk geïnteresseerd zijn in mogelijke veranderingen tussen

2003 en 2011, nemen we ook het afnamejaar op in de analyses als verklarende variabele. We zien hierbij dat de proportie variantie erg gelijk verdeeld blijft over de verschillende niveaus.

Tabel 10. Proportie variantie op school-, klas- en leerlingniveau in 2003 en 2011 tesamen, leeg model en model met controle voor afnamejaar.

	N	2003+2011: leeg model			2003+2011: controle afnamejaar		
		Variantie	S.E.	ICC (%)	Variantie	S.E.	ICC (%)
Schoolniveau	291	467.04	69.59	13.24	466.73	70.16	13.23
Klasniveau	526	142.80	41.13	4.05	142.85	41.16	4.05
Leerlingniveau	9351	2917.13	55.62	82.71	2917.13	55.61	82.72

In Tabel 11 (zie p.31) staan de resultaten van de analyses weergegeven. In Model 0 wordt enkel afnamejaar toegevoegd. De erg lichte daling in gemiddelde wiskundeprestatie in Vlaanderen (551 in 2003 ten opzichte van 550 in 2011) is niet significant. In Model 1 voegen we, analoog aan de analyses in de voorgaande paragraaf, geslacht, geboortemaand en al dan niet normaalvorderend zijn toe als controlevariabelen. In lijn met de afzonderlijke analyses voor 2003 en 2011, zien we dat geslacht en al dan niet normaalvorderend significant positief samenhangen met de wiskundeprestaties. Geboortemaand vertoont een significante negatieve samenhang met de wiskundeprestaties (analoog aan de resultaten voor 2011 afzonderlijk). In Model 2A voegen we de variabelen aantal boeken thuis en thuistaal van de leerlingen toe aan het model. Analooog aan de analyses afzonderlijk voor 2003 en 2011, bestaat er een significant positieve samenhang van aantal boeken thuis met de wiskundeprestaties en een significant negatieve samenhang tussen een andere thuistaal van de leerling en diens wiskundeprestaties.

Om uitspraken te kunnen doen over de mate waarin de samenhang tussen het aantal boeken thuis en thuistaal enerzijds en wiskundeprestaties anderzijds, gewijzigd is tussen 2003 en 2011, voegen we de interactietermen van deze variabelen en afnamejaar toe in Model 2B. Hiermee gaan we na of de samenhang tussen het aantal boeken thuis en thuistaal en prestaties verschilt in 2003 ten opzichte van 2011. Aangezien de interactietermen niet significant zijn, betekent dit dat de samenhang in 2003 en in 2011 dezelfde is. Grafiek 11 en Grafiek 12 (zie p.31) tonen hoe deze relaties verlopen, zowel in 2003 als in 2011. Hoewel de regressielijnen voor 2011 niet significant verschillen van die voor 2003, zien we wel dat deze lijnen lichtjes vlakker lopen in 2011 dan in 2003. Dit wijst in de richting van een afname van de samenhang.⁹ Dit is echter een niet-significante en erg minieme tendens.¹⁰

⁹ Idealiter zouden we een regressielijn van prestaties op aantal boeken thuis en thuistaal willen bekomen die horizontaal ligt, met een hoog intercept. Een hoog intercept betekent dat hoge prestaties bekomen worden door respectievelijk

Aangezien deze interactietermen niet significant zijn, en de fit van het model (test op basis van Log Likelihood index) aangeeft dat het model geen beter model is dan wanneer de interactietermen niet worden opgenomen (Model 2A), besluiten we om deze interactietermen niet te behouden in het model. In een volgend model, Model 3A, voegen we de schoolcompositiekenmerken toe, met name het gemiddeld aantal boeken thuis en het percentage anderstaligen. Ook hier lopen de resultaten gelijk met de resultaten waarbij we 2003 en 2011 afzonderlijk bekijken: er is een significante positieve samenhang tussen schoolcompositie in termen van SES en wiskundeprestaties en een significante negatieve samenhang tussen het percentage anderstaligen op school en wiskundeprestaties. Zoals aangegeven, is er een erg minieme, niet-significante achteruitgang in wiskundeprestaties in 2011 in vergelijking met 2003. Wanneer echter gecontroleerd wordt voor de schoolcompositie, bekomen we bij het afnamejaar een positieve coëfficiënt. Dit betekent dat, na controle voor leerlingkenmerken en de schoolcompositie inzake SES en thuistaal, de wiskundeprestaties in 2011, hoewel erg miniem en niet-significant, hoger liggen dan deze in 2003.

Ook hier zijn we erin geïnteresseerd of de samenhang tussen (1) schoolcompositie inzake SES en wiskundeprestaties en (2) schoolcompositie inzake thuistaal en wiskundeprestaties verschillend is voor 2011 en voor 2003. Daarom voegen we in Model 3B de interactietermen toe tussen afnamejaar en de schoolcompositievariabelen. Aangezien deze niet significant zijn, betekent dit dat de samenhang tussen deze schoolcompositiekenmerken en wiskundeprestaties niet significant gewijzigd zijn in 2011 ten opzichte van 2003. In

Grafiek 13 en Grafiek 14 (zie p.33) visualiseren we deze niet-significante interacties. Analoog aan de interacties tussen het afnamejaar en de leerlingvariabelen aantal boeken thuis en thuistaal, zien we

leerlingen met weinig boeken thuis en leerlingen die thuis steeds Nederlands spreken, en de horizontale ligging zou betekenen dat leerlingen met een ander aantal boeken thuis en met een andere thuistaal, hetzelfde hoog prestatieniveau behalen. Dit betekent dat er geen samenhang bestaat tussen aantal boeken thuis en thuistaal aan de ene kant en prestaties aan de andere kant.

¹⁰ Om een zinvolle interpretatie aan het interactie-effect te geven, is het nodig om de controlevariabelen op leerlingniveau te centreren rond hun algemeen gemiddelde. Op deze manier zien we in de interactieplot hoe het effect van de variabele (met name respectievelijk aantal boeken thuis, thuistaal, schoolcompositie inzake SES en schoolcompositie in termen van thuistaal) is voor een 'gemiddelde leerling', afzonderlijk voor 2003 en 2011. Anders gezegd: voor een leerling met een 'gemiddeld' geslacht, een gemiddelde geboortemaand, een 'gemiddelde' m.b.t. normaalvorderend zijn, het gemiddelde aantal boeken thuis en de gemiddelde thuistaal. Dit gebeurt door de score van elke leerling te verminderen met het gemiddelde. Bv: het centreren van de geboortemaand gebeurt door bij elke leerling de score op geboortemaand (0-11) te verminderen met het gemiddelde over 2003 en 2011. Omdat het niet zinvol is om uitspraken te doen over de wiskundeprestaties van een leerling met een 'gemiddeld' geslacht, maar wel over de wiskundeprestaties van jongens versus meisjes, hebben we dit centreren enkel doorgevoerd om de interacties te plotten. Hiervoor werden een aantal aparte analyses gedaan, waarbij de variabelen op leerlingniveau gecentreerd opgenomen werden in de analyses. Hierdoor wijkt het intercept mogelijks af van het intercept in Tabel 10. Wel zorgt dit ervoor dat de interacties interpreteerbaar zijn, met name zien we in de interactieplots de invloed van de verschillende variabelen (respectievelijk aantal boeken thuis, thuistaal, schoolcompositie in termen van SES en schoolcompositie inzake thuistaal), apart voor 2003 en 2011, telkens voor het gemiddelde leerlingenpubliek (in geslacht, geboortemaand, normaalvorderend zijn, aantal boeken thuis en/of thuistaal).

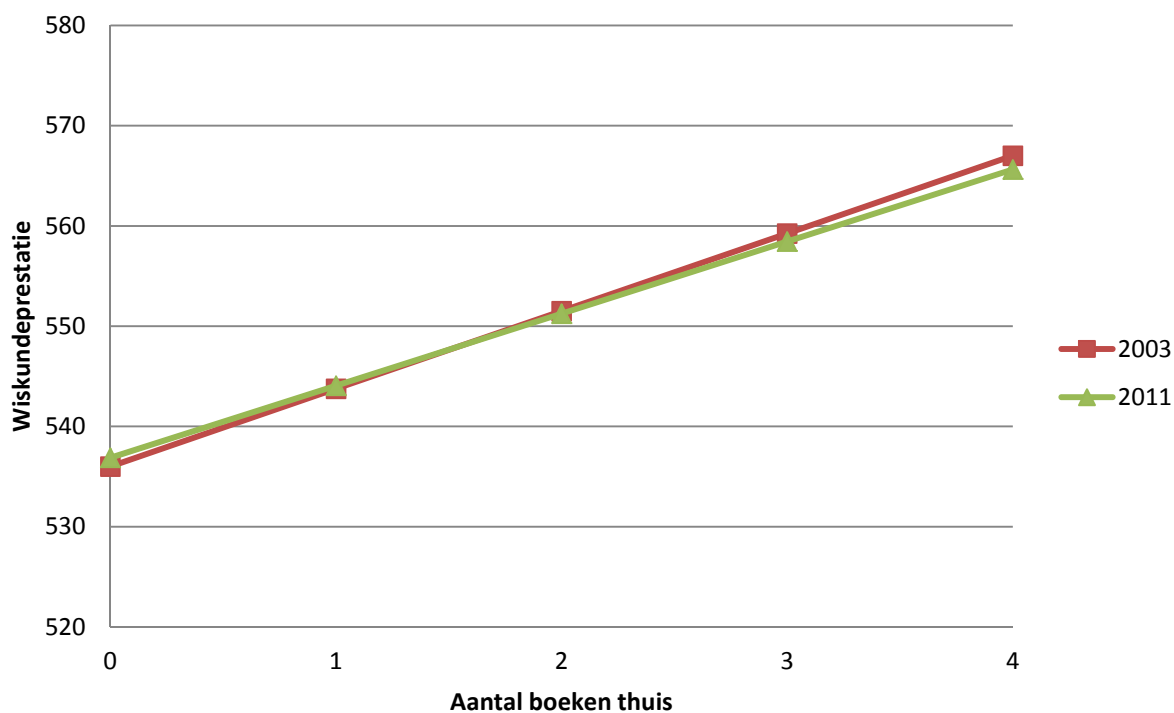
hier dat de regressielijnen in 2011 een minder steile helling vertonen dan in 2003 en dus dat er een trend waar te nemen is waarbij de invloed van de schoolcompositie inzake SES en taal een minder grote samenhang vertoont met wiskundeprestaties in 2011 dan in 2003. We vermelden nogmaals, dat deze trend echter niet significant is.

Tabel 11. Invloed aantal boeken thuis, thuistaal en schoolcompositie op wiskundeprestaties: veranderingen tussen 2003 en 2011.

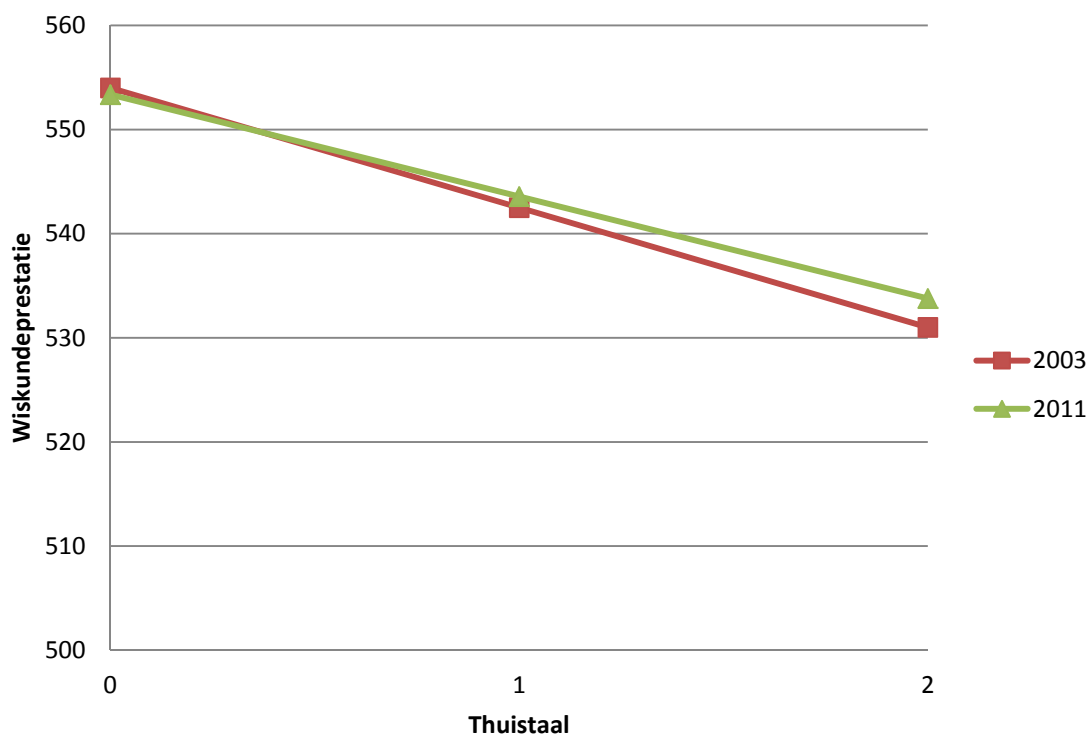
	Model 0		Model 1		Model 2A		Model 2B		Model 3A		Model 3B	
	b	S.E.	b	S.E.	b	S.E.	b	S.E.	b	S.E.	b	S.E.
Intercept	551.36***	1.47	558.02***	1.91	544.14***	2.46	543.76***	2.78	536.91***	6.17	534.84***	8.71
Afnamejaar ^a	-.78	2.22	-1.39	2.05	-.23	1.98	0.46	3.55	3.37	1.93	8.26	12.51
<i>Leerlingniveau</i>												
Geslacht			6.35***	1.22	6.77***	1.21	6.79***	1.20	6.79***	1.20	6.82***	1.20
Geboortemaand			-.51*	.21	-.47*	.20	-.47*	.20	-.48*	.20	-.48*	.20
Normaalvorderend			40.43***	1.62	36.47***	1.55	36.47***	1.55	36.03***	1.55	36.03***	1.55
Aantal boeken thuis					7.49***	.78	7.75***	.98	7.27***	.80	7.27***	.80
Thuistaal					-10.63***	1.40	-11.50***	2.17	-8.88***	1.41	-8.86***	1.41
<i>Interactietermen</i>												
Afnamejaar*boeken							-0.56	1.26				
Afnamejaar*thuistaal							1.71	3.21				
<i>Schoolniveau</i>												
Compositie SES									6.72*	2.89	8.30	4.16
Compositie thuistaal									-38.52***	4.91	-45.42***	6.63
<i>Interactietermen</i>												
Afnamejaar*CompSES											-3.51	6.00
Afnamejaar*CompTaal											10.53	9.20
<i>Variantie (random gedeelte)</i>												
School (intercept)	466.73	70.16	309.28	50.68	250.07	45.82	250.29	45.61	174.42	39.76	173.48	39.65
Klas (intercept)	142.85	41.16	137.68	33.65	134.33	33.85	134.33	33.83	143.12	35.39	141.86	34.96
Leerling	2917.13	55.61	2633.99	53.09	2562.68	54.52	2562.32	54.11	2561.46	54.49	2561.65	54.51
Totale proportie verklaarde variantie			12.64		16.44		16.44		18.37		18.42	
Log Likelihood	-50929.30		-50426.00***		-50281.1***		-50280.60		-50255.6***		-50254.7	

* p < .05; ** p < .01; *** p < .001; ^a 0 = afnamejaar 2003; 1 = afnamejaar 2011

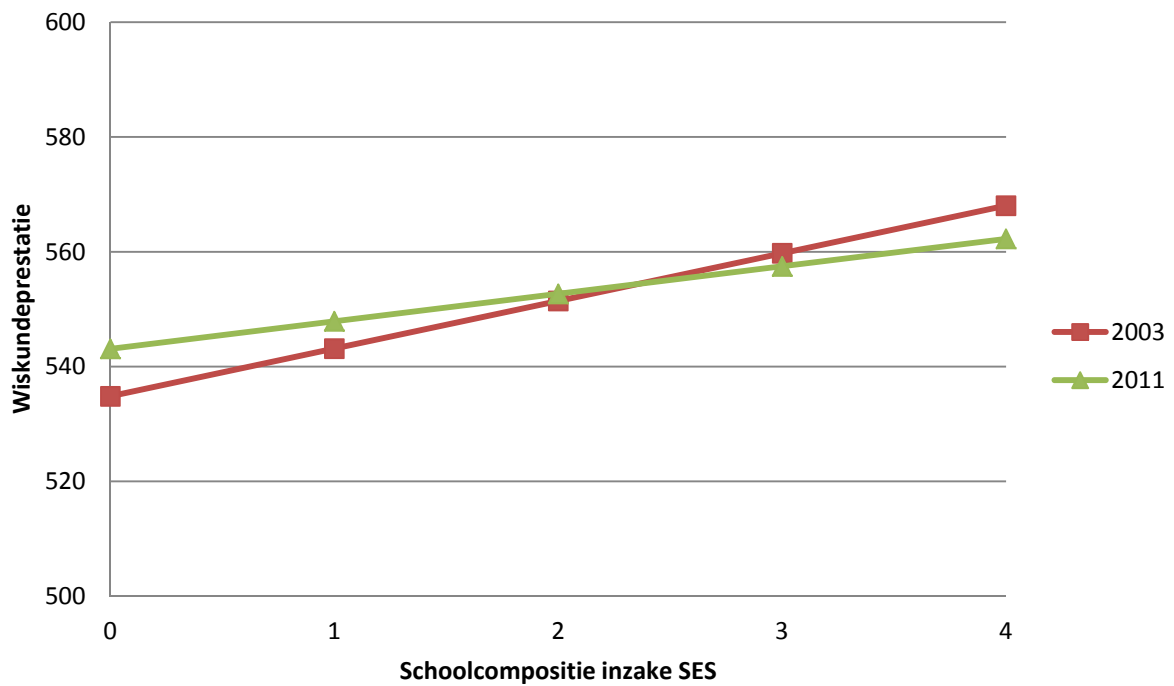
Grafiek 11. Niet-significant interactie-effect van aantal boeken thuis en afnamejaar op wiskundeprestaties.



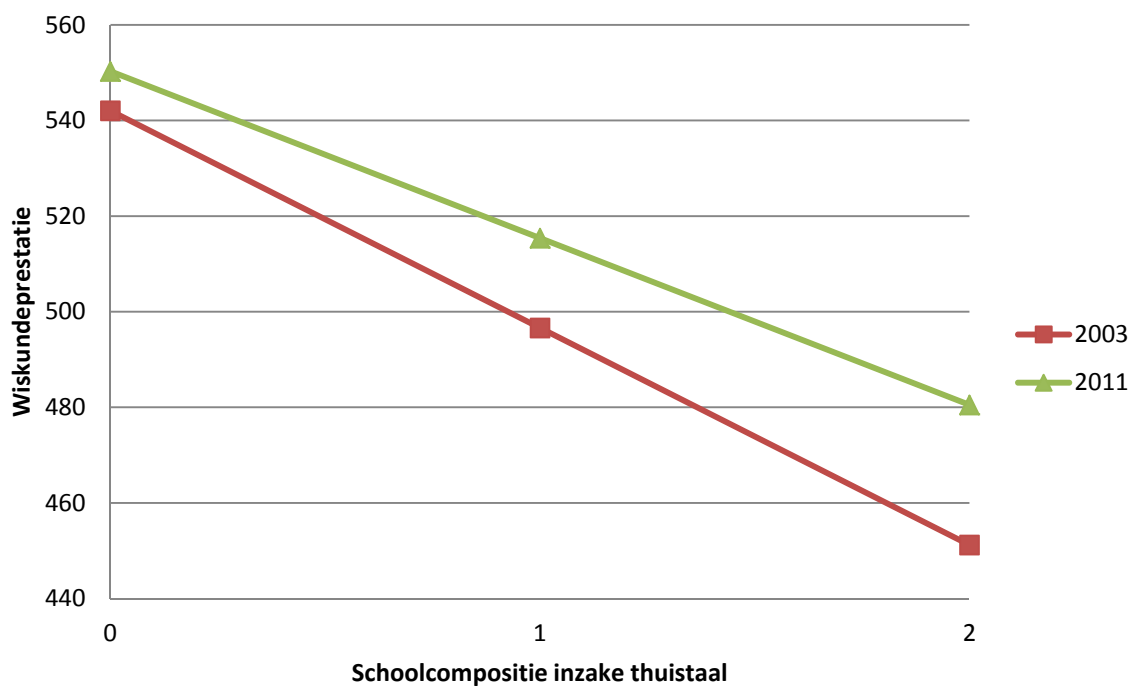
Grafiek 12. Niet-significant interactie-effect van thuistaal en afnamejaar op wiskundeprestaties.



Grafiek 13. Niet-significant interactie-effect van afnamejaar en schoolcompositie inzake SES op wiskundeprestaties.



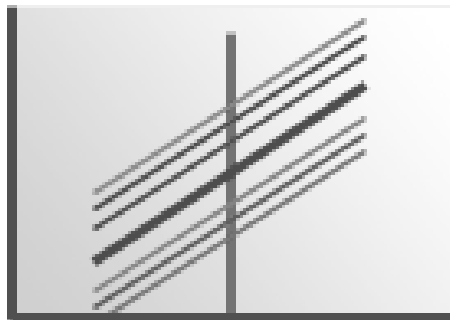
Grafiek 14. Niet-significant interactie-effect van afnamejaar en schoolcompositie inzake thuistaal op wiskundeprestaties.



3.2. Verschillen tussen scholen in Vlaanderen in sociale ongelijkheid en ongelijkheid op basis van thuistaal in wiskunde

In de bovenstaande analyses werd er vanuit gegaan dat de invloed van het aantal boeken thuis, de thuistaal en de schoolcompositie inzake aantal boeken thuis en thuistaal, in elke school op een zelfde wijze samenhangt met de prestaties. Met name werd er verondersteld dat het intercept verschillend kan zijn in van school tot school (random intercept), maar dat de helling van de relatie voor elke school dezelfde is (fixed slope). Dit wordt visueel voorgesteld in Grafiek 15. Er werd verondersteld dat met het stijgen van het aantal boeken thuis met 1 unit, de prestaties in elke school eenzelfde stijging kennen; of dat het stijgen van de score van thuistaal met 1 unit (bv: gaande van altijd Nederlands thuis naar soms Nederlands thuis) in elke school zorgt voor eenzelfde daling van de wiskundeprestaties.

Grafiek 15. Random intercept, fixed slope



Er kan echter aangenomen worden dat scholen verschillen in de mate waarin deze kenmerken een samenhang vertonen met de wiskundeprestaties. Of anders gezegd: het kan zijn dat de invloed van het aantal boeken thuis en de thuistaal van de leerling anders is in de ene school dan in de andere school. Of nog anders: er kan verondersteld worden dat de sociale ongelijkheid en/of ongelijkheid o.b.v. thuistaal groter is in de ene school dan in de andere school. Dit wordt visueel voorgesteld in Grafiek 16. In deze grafiek is te zien dat niet enkel het intercept kan verschillen tussen scholen (random intercept), maar dat ook de helling/slope van de regressielijn verschillend is tussen scholen (random slope). In de uiterst linkse voorstelling bestaat er een positieve relatie tussen het intercept en de slope van de regressielijn; hoe hoger het intercept, hoe groter de slope. In de middelste voorstelling bestaat er een negatieve relatie tussen het intercept en de slope van de regressielijn; hoe hoger het intercept, hoe minder steil de slope. En in de laatste voorstelling bestaat er geen

verband tussen het intercept en de slope van de regressielijn; sommige scholen hebben een laag/hoog intercept en een grote slope, andere scholen hebben een laag/hoog intercept en een kleine slope.

Grafiek 16. Random intercept en random slope, achtereenvolgens een positieve, een negatieve en geen samenhang tussen intercept en helling.



In dit deel wordt dan ook nagegaan in welke mate scholen verschillen wat betreft sociale ongelijkheid en ongelijkheid o.b.v. thuistaal in wiskundeprestaties. Met name gaan we eerst na in welke mate er verschillen zijn tussen scholen betreffende sociale ongelijkheid en ongelijkheid o.b.v. thuistaal in wiskundeprestaties, respectievelijk in 2003 en in 2011. In tweede instantie bekijken we ook hier in welke mate deze schoolverschillen gewijzigd zijn in 2011 in vergelijking met 2003.

3.2.1. Verschillen tussen scholen in de invloed van leerlingkenmerken en schoolcompositie op wiskundeprestaties in 2003 en 2011

In deze paragraaf staan volgende onderzoeksvragen centraal:

- Verschillen scholen in de samenhang tussen het aantal boeken thuis en wiskundeprestaties, respectievelijk in 2003 en 2011, na controle voor geslacht, geboortemaand, al dan niet normaalvorderend zijn, thuistaal, schoolcompositie inzake SES en percentage anderstaligen op school? En zo ja, hoe verschillen scholen?
- Verschillen scholen in de samenhang tussen thuistaal en wiskundeprestaties, respectievelijk in 2003 en 2011, na controle voor geslacht, geboortemaand, al dan niet normaalvorderend zijn, aantal boeken thuis, schoolcompositie inzake SES en percentage anderstaligen op school? En zo ja, hoe verschillen scholen?

We starten met het model dat de data het best fit en waarmee we in paragraaf 3.1.1 geëindigd zijn, met name het model waarbij geslacht, geboortemaand, aantal boeken thuis, thuistaal, schoolcompositie in termen van SES en percentage anderstaligen op school als verklarende variabelen

genomen worden om de wiskundeprestaties van leerlingen te schatten. De resultaten hiervan staan weergegeven in Model 3 (Tabel 13), respectievelijk voor de data van 2003 en 2011. In dit model werd enkel verondersteld, zoals eerder gezegd, dat scholen kunnen verschillen in intercept, maar dat de samenhang tussen de leerlingkenmerken en de prestaties in elke school dezelfde is (fixed slope). Om na te gaan of er verschillen zijn tussen scholen wordt in de analyses opgenomen dat de slope/helling van de regressielijn van wiskundeprestaties op aantal boeken thuis kan variëren tussen scholen. Tabel 13 toont de resultaten van de modellen, waarbij achtereenvolgens aantal boeken thuis (Model 4A) en thuistaal (Model 4B) als random slope in het model werden opgenomen, respectievelijk voor 2003 en 2011. Om vervolgens een antwoord te bieden op de onderzoeksvragen, dient gekeken te worden of het model waarbij random slopes worden geïntroduceerd, een betere fit met de data vertoont dan het model waarbij enkel een random intercept verondersteld wordt en een fixed slope (zie illustratie in Grafiek 15). In Tabel 12 staat door middel van de likelihood ratio test aangegeven of het model met random slopes een significant beter model is dan het model waarbij geen random slopes verondersteld worden. In deze tabel is te zien dat in 2003, zowel het model waarbij een random slope tussen scholen verondersteld wordt met betrekking tot het aantal boeken thuis en wiskundeprestaties (Model 4A) als het model waarbij een random slope tussen scholen verondersteld wordt met betrekking tot de thuistaal en wiskundeprestaties (model 4B) significant betere modellen zijn dan het model waarbij enkel een random intercept verondersteld wordt (Model 3). Dit betekent dat in 2003 scholen verschillen in de relatie tussen enerzijds aantal boeken thuis en wiskundeprestaties en anderzijds thuistaal en wiskundeprestaties. In 2011 zien we eenzelfde resultaat voor aantal boeken thuis: in 2011 verschillen scholen in de relatie tussen het aantal boeken thuis en wiskundeprestaties. Model 4B is niet significant, wat betekent dat het model waarbij enkel een random intercept verondersteld wordt, de data het best fit of dat m.a.w. aangenomen kan worden dat de relatie tussen thuistaal en wiskundeprestaties in elke school dezelfde is.

Tabel 12. Model fit van random slope op schoolniveau en correlatie tussen slope en intercept, achtereenvolgens voor aantal boeken thuis en thuistaal, respectievelijk in 2003 en 2011.

	2003					2011				
	Log Likelihood	Chi-square	Df	p	R Slope-intercept	Log Likelihood	Chi-square	Df	p	R slope-intercept
Model 3	-24623.2					-25629.0				
Model 4A: Boeken	-24618.9	8.54	2	.000	-0.73	-25623.7	10.61	2	.000	-0.81
Model 4B: Thuistaal	-24615.2	15.91	2	.000	0.00	-25627.7	2.62	2	.270	0.08

Tabel 13. Model 3 in 2003 en 2011, waarbij achtereenvolgens random slopes op schoolniveau worden geïntroduceerd voor aantal boeken thuis (Model 4A) en thuistaal (Model 4B).

	2003						2011					
	Model 3		Model 4A		Model 4B		Model 3		Model 4A		Model 4B	
	b	S.E.	b	S.E.	b	S.E.	b	S.E.	b	S.E.	b	S.E.
Intercept	534.59 ^{***}	8.77	534.38 ^{***}	8.49	539.24 ^{***}	9.07	543.21 ^{***}	8.39	542.89 ^{***}	8.49	544.88 ^{***}	9.18
<i>Leerlingniveau</i>												
Geslacht ^a	5.70 ^{**}	1.74	5.72 ^{**}	1.70	5.58 ^{**}	1.71	8.06 ^{***}	2.15	8.04 ^{***}	2.15	8.07 ^{**}	2.18
Geboortemaand ^b	-.25	.26	-0.24	0.26	-0.29	0.26	-.73	.37	-0.75 [*]	0.36	-0.74	0.36
Normaalvorderend ^c	37.28 ^{***}	2.27	37.24 ^{***}	2.28	36.72 ^{***}	2.19	34.77 ^{***}	2.19	34.90 ^{***}	2.17	34.72 ^{***}	2.19
Aantal boeken thuis ^d	7.45 ^{***}	1.03	7.33 ^{***}	0.87	7.38 ^{***}	1.04	7.04 ^{***}	1.04	6.98 ^{***}	0.91	7.02 ^{***}	1.04
Thuistaal ^e	-9.68 ^{***}	2.19	-9.54 ^{***}	2.17	-12.29 ^{***}	1.78	-7.99 ^{***}	2.19	-8.11 ^{***}	2.15	-8.05 ^{***}	1.96
<i>Schoolniveau</i>												
Compositie SES ^f	7.94	4.20	8.03	4.03	6.18	4.22	5.32	4.27	5.29	4.23	4.64	4.47
Compositie thuistaal ^g	-43.94 ^{***}	6.56	-42.30 ^{***}	6.58	-48.79 ^{***}	7.11	-36.31 ^{***}	6.99	-33.48 ^{***}	7.15	-37.30 ^{***}	7.45
<i>Variantie (random gedeelte)</i>												
Leerling (intercept)	2546.01	72.21	2518.60	75.77	2516.61	71.40	2573.87	77.07	2548.50	79.12	2563.49	75.49
Klas (intercept)	120.28	43.21	120.14	42.19	120.27	44.08	168.54	57.03	175.68	62.56	167.80	58.25
<i>School</i>												
Intercept	166.43	40.04	326.95	103.13	153.76	43.82	176.87	74.12	381.00	155.28	167.19	67.09
Slope aantal boeken			26.80	19.76					24.58	15.16		
Slope thuistaal					146.86	65.68					52.02	74.04

* p < .05; ** p < .01; *** p < .001

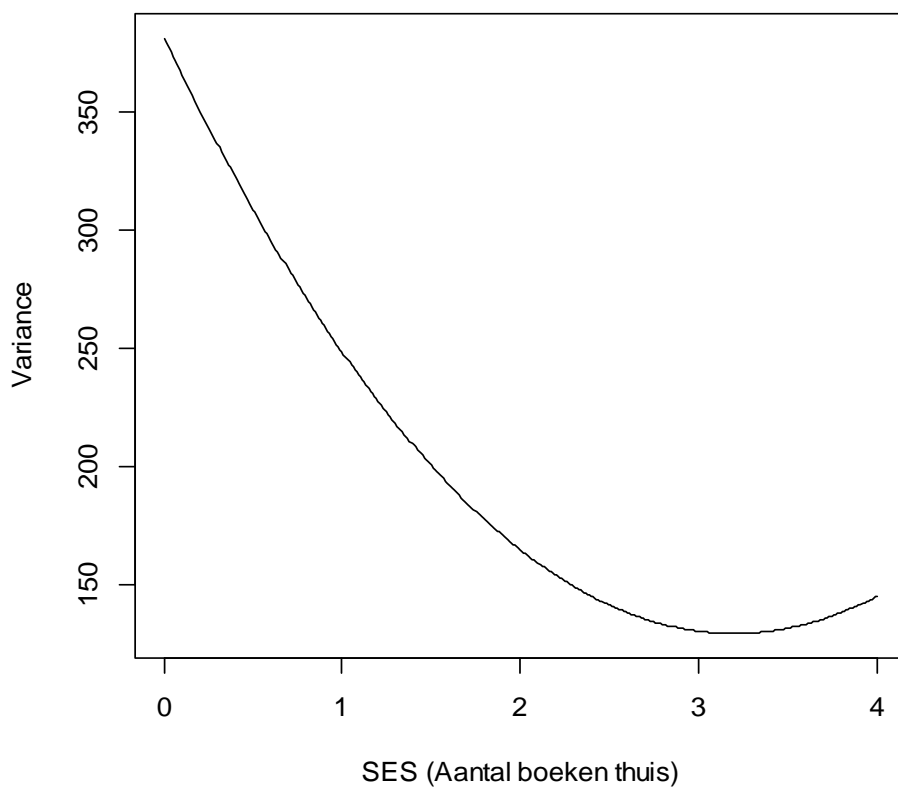
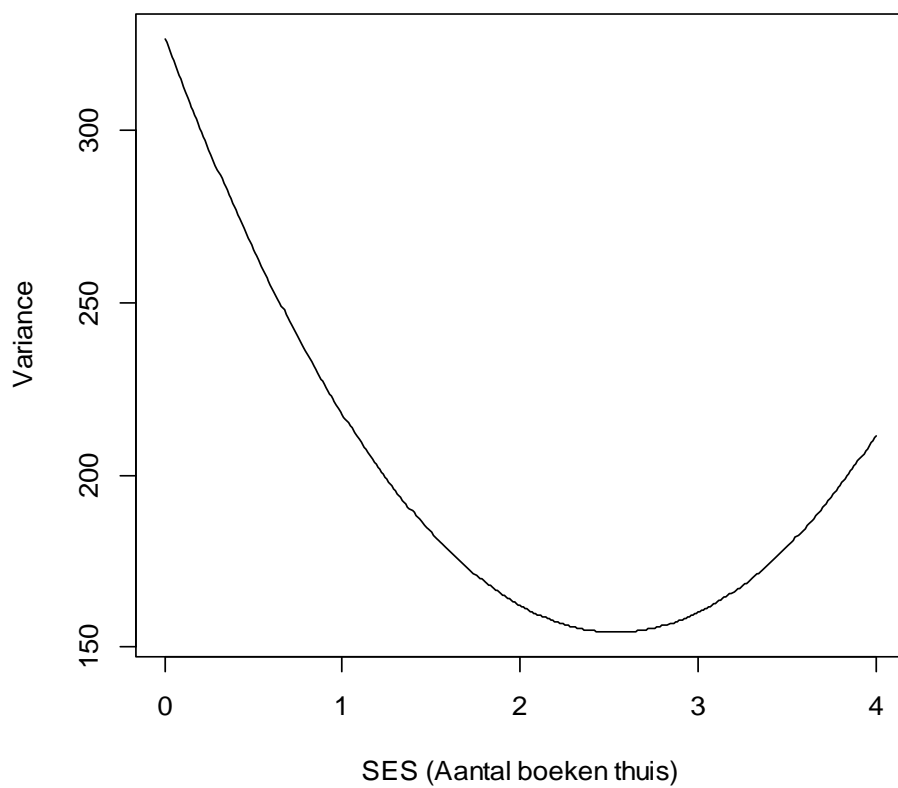
Vervolgens willen we weten hoe scholen verschillen in de relatie tussen aantal boeken thuis (2003 en 2011) en thuistaal (enkel in 2003, aangezien het model in 2011 niet significant beter is dan het model met fixed slopes). Hiervoor kijken we naar de relatie tussen het intercept en de slope (R intercept-slope). Wat betreft de random slope voor aantal boeken thuis, zien we dat de correlatie tussen het intercept en de slope zowel in 2003 als in 2011 negatief is. Dit betekent dat een hoog intercept samenhangt met een kleine slope en omgekeerd, dat een laag intercept samenhangt met een hoge slope.

Om een beter beeld te krijgen van de verschillen tussen scholen, plotten we de variantie tussen scholen in prestaties in relatie met het aantal boeken thuis (2003 en 2011) (zie Grafiek 17). Wat we hierbij zien, is dat er enorm grote verschillen zijn tussen scholen wat betreft de prestaties van leerlingen met een lage SES. De verschillen tussen scholen verkleinen naarmate leerlingen een hogere SES hebben (of meer boeken thuis hebben). In 2003 zien we de kleinste verschillen tussen scholen bij leerlingen met een SES van ca. 2.5 (d.w.z. ca. 60 à 150 boeken thuis); in 2011 ligt het kleinste verschil tussen scholen bij leerlingen met een SES van ca. 3 (d.w.z. tussen 100 en 200 boeken thuis). Nadien neemt de variantie in wiskundeprestaties terug toe, d.w.z. de verschillen tussen scholen zijn opnieuw groter bij leerlingen met een SES hoger dan respectievelijk 2.5 en 3.

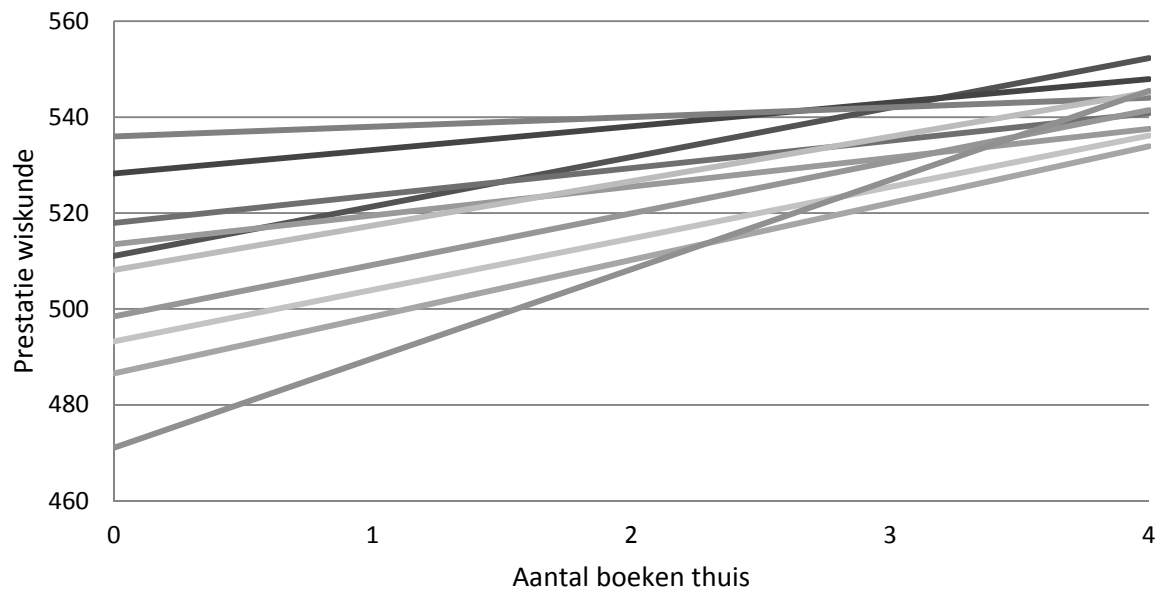
Aangezien in beide afnamejaren ca. 150 scholen deelnamen aan het onderzoek, is een grafiek die de regressielijn toont van het aantal boeken thuis op prestaties voor elke school afzonderlijk, niet overzichtelijk. Om toch een illustratie te geven van deze relatie en van de verschillen tussen scholen in de invloed van aantal boeken thuis op wiskundeprestaties, werd Grafiek 18 toegevoegd, waarin illustratief de verschillen tussen scholen visueel worden voorgesteld aan de hand van een aantal regressies voor een aantal scholen. Hierbij zien we dat scholen met een hoger intercept een vlakker regressielijn hebben dan scholen met een lager intercept; of nog: scholen die ook bij de leerlingen met een lage SES hoge prestaties behalen, slagen er beter in om de samenhang tussen het aantal boeken thuis en wiskundeprestaties zo klein mogelijk te houden. Daarbij zien we ook dat alle lijnen convergeren bij een waarde drie en dat ze bij waarden >3 opnieuw wat uitwaaien. Dit is de representatie van de variantieplot, waarbij in het dieptepunt van 2.5 en 3 de minst grote variantie tussen scholen waar te nemen is, ofwel: waarbij scholen zo goed als dezelfde prestaties behalen bij leerlingen met die bijbehorende SES/aantal boeken thuis. Verder betekent dit ook dat scholen differentieel effectief zijn voor leerlingen met een verschillende SES. Een leerling met een hoge SES (aantal boeken thuis waarde 4, dit is meer dan 200 boeken thuis) kan best les volgen in een school met een laag intercept (ofwel een school die lage prestaties behaalt bij leerlingen met een lage SES). Een leerling met een lage SES (aantal boeken thuis waarde 0, dit is minder dan 10 boeken thuis) en

gemiddelde SES (tot de waarde van aantal boeken thuis 2.5 à 3, dit is tussen 60 en 200 boeken thuis) kan best onderwijs volgen in een school met een hoog intercept. Voor leerlingen met een score 2 á 3 betreffende aantal boeken thuis (d.w.z. leerlingen met 100 tot 200 boeken thuis), is de schoolkeuze van minder belang, aangezien de verschillen tussen scholen in wiskundeprestatie voor hen erg klein zijn.

Grafiek 17. Plot voor de variantie in prestatie tussen scholen in relatie met aantal boeken thuis in 2003 (boven) en 2011 (onder).

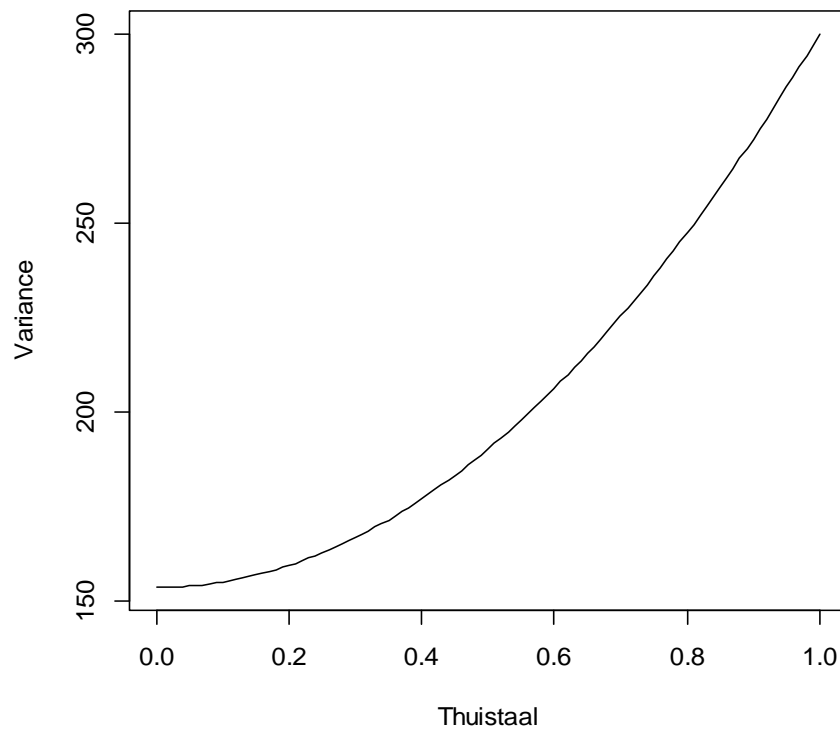


Grafiek 18. Illustratie van de verschillen tussen scholen in 2003: random intercept, random slope en negatieve relatie tussen intercept en slope, met kantelpunt rond waarde 3 (dit is 101-200 boeken thuis).

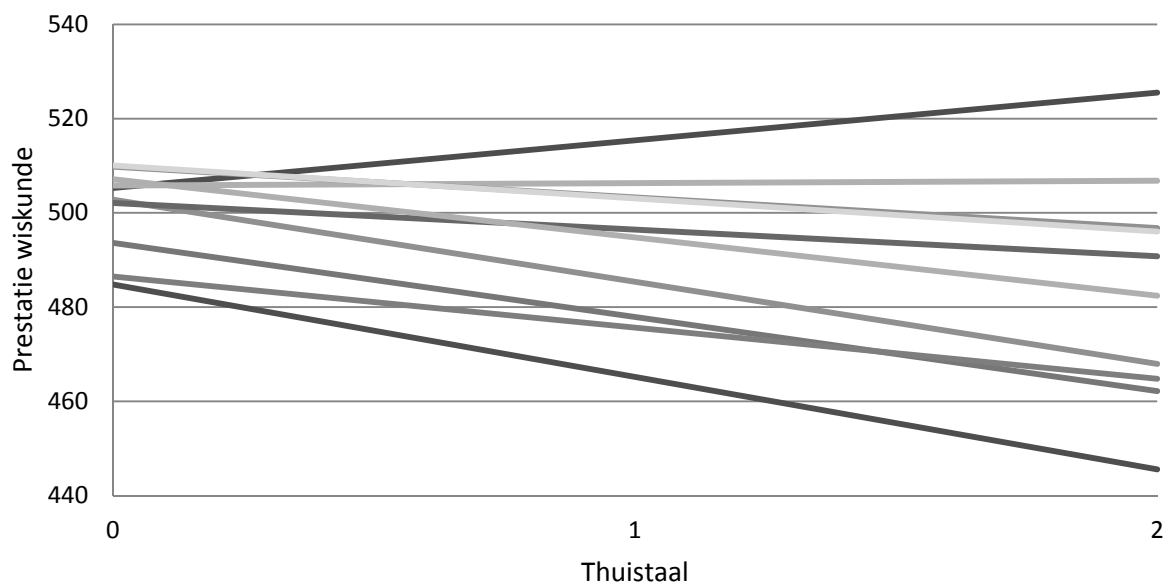


In 2003 is ook de samenhang tussen thuistaal en wiskundeprestaties verschillend tussen scholen. De slope van de regressielijnen is negatief (Model 4B, Tabel 13), wat betekent dat scholen hogere prestaties behalen voor leerlingen die altijd Nederlands spreken dan voor leerlingen die soms Nederlands spreken thuis en dat op hun beurt leerlingen die soms Nederlands spreken thuis betere prestaties behalen dan leerlingen die thuis nooit Nederlands spreken. Bekijken we ook hier de variantie tussen scholen (zie Grafiek 19), dan zien we dat er weinig variantie is tussen scholen voor leerlingen die altijd Nederlands spreken thuis (thuistaal = 0), en een toenemende variantie tussen scholen voor leerlingen die soms en nooit Nederlands spreken thuis. Dit betekent dat, in 2003, voornamelijk voor leerlingen die soms en nooit Nederlands spreken thuis, er grote verschillen bestaan tussen scholen in de invloed van de thuistaal op prestaties. De correlatie tussen het intercept en de slope is .00 (zie Tabel 12), wat betekent dat er geen samenhang bestaat tussen het intercept en de slope van de regressie van wiskundeprestaties op thuistaal in 2003. Dit komt overeen met de rechtse illustratie in Grafiek 16. Aangezien er geen samenhang is, zullen verschillende scholen een verschillende regressielijn hebben, wat betekent dat er scholen kunnen zijn met een hoog intercept en een kleine (negatieve) slope, scholen met een hoog intercept en een grote (negatieve) slope, scholen met een laag intercept en een kleine (negatieve) slope en scholen met een laag intercept en grote (negatieve) slope. Dit wordt geïllustreerd in Grafiek 20.

Grafiek 19. Plot voor de schoolvariantie inzake thuistaal in 2003.



Grafiek 20. Illustratie van de verschillen tussen scholen in 2003 in thuistaal: random intercept, random slope en geen relatie tussen intercept en slope.



3.2.2. Veranderingen tussen 2003 en 2011 in de verschillen tussen scholen met betrekking tot de invloed van achtergrondkenmerken op wiskundeprestaties

In paragraaf 3.2.1 zagen we dat er verschillen tussen scholen bestaan in de samenhang tussen het aantal boeken thuis en wiskundeprestaties, zowel in 2003 als in 2011. Zo zagen we dat in sommige scholen de samenhang tussen aantal boeken thuis en wiskundeprestaties groter is dan in andere scholen. Met betrekking tot de thuistaal zagen we dat er in 2003 verschillen bestaan tussen scholen betreffende de samenhang tussen thuistaal en wiskundeprestaties, terwijl in 2011 geen verschillen meer tussen scholen te merken zijn hieromtrent. Op basis van de analyses in voorgaande paragraaf, kunnen echter geen uitspraken gedaan worden over de vraag of deze verschillen tussen scholen significant zijn gewijzigd (toegenomen of afgenomen) tussen 2003 en 2011. Dit gaan we na in deze laatste paragraaf. Met name bieden we een antwoord op volgende twee onderzoeksvragen:

- Zijn de verschillen tussen scholen in de samenhang tussen het aantal boeken thuis en wiskundeprestaties gewijzigd tussen 2003 en 2011, na controle voor geslacht, geboortemaand, al dan niet normaalvorderend zijn, thuistaal, schoolcompositie inzake SES en percentage anderstaligen op school?
- Zijn de verschillen tussen scholen in de samenhang tussen de thuistaal en wiskundeprestaties gewijzigd tussen 2003 en 2011, na controle voor geslacht, geboortemaand, al dan niet normaalvorderend zijn, aantal boeken thuis, schoolcompositie inzake SES en percentage anderstaligen op school?

Analoog aan de analyses in paragraaf 3.1.2, bekijken we in deze analyses de data van 2003 en 2011 gezamenlijk in één dataset en nemen we afnamejaar op als een verklarende variabele. We vertrekken van Model 3A (zie Tabel 11), waarbij geslacht, geboortemaand en al dan niet normaalvorderend als controlevariabelen worden opgenomen, het aantal boeken thuis en thuistaal op leerlingenniveau als indicatoren van SES en etniciteit en de schoolcompositie inzake SES en thuistaal als variabelen op schoolniveau worden opgenomen ter verklaring van de wiskundeprestaties. In Model 3B (Tabel 15) bekijken we hetzelfde model, maar met een random slope voor aantal boeken thuis. Met name bekijken we over beide afnamejaren heen in welke mate de samenhang tussen het aantal boeken thuis en wiskundeprestaties verschilt tussen scholen. In Tabel 14 zien we dat dit model de data beter fit dan het model waarbij enkel het intercept kan variëren tussen scholen. Dit betekent dat scholen verschillen in samenhang tussen aantal boeken thuis en wiskundeprestaties, bekeken met de data van 2003 en 2011 tesamen. In paragraaf 3.1.2

gingen we na in welke mate het aantal boeken thuis een andere samenhang vertoont met wiskundeprestaties door middel van de interactieterm 'aantal boeken thuis * afnamejaar'. Deze interactieterm was niet significant. Om vervolgens uitspraken te doen over de mate waarin de schoolverschillen in samenhang tussen aantal boeken thuis en wiskundeprestaties, verschillen in 2011 ten opzichte van 2003, voegen we deze interactieterm opnieuw toe in Model 4A. Analoog aan de analyses in paragraaf 3.1.2., zien we dat de opname van deze interactieterm niet significant is en niet zorgt voor een betere fit van de data. Wanneer we vervolgens ook een random slope van deze interactieterm toevoegen in Model 4B, gaan we na in welke mate de verschillen tussen scholen in samenhang tussen aantal boeken thuis en wiskundeprestaties gewijzigd zijn in 2011 ten opzichte van 2003. Tabel 14 toont de fitmaten van Model 4B ten opzichte van Model 4A, waarbij te zien is dat het eerstgenoemde model geen beter model is dan Model 4A. Op basis hiervan besluiten we dat de verschillen tussen scholen in de samenhang tussen aantal boeken thuis en wiskundeprestaties niet significant gewijzigd zijn in 2011 ten opzichte van 2003.

Tabel 14. Model fit van random slopes op schoolniveau, voor aantal boeken thuis en interactieterm 'aantal boeken thuis * afnamejaar'.

	2003-2011			
	Log Likelihood	Chi-square	Df	p
Model 3A	-50255.6			
Model 3B: Random slope Aantal boeken thuis	-50246.2	18.79	2	.000
Model 4A: toevoeging interactieterm 'boeken*afnamejaar'	-50246.0	.40	1	.527
Model 4B: Random slope interactieterm	-50245.5	1.00	7	.995

Tabel 15. Invloed aantal boeken thuis, thuistaal en schoolcompositie op wiskundeprestaties (Model 3A), met een random slope voor aantal boeken thuis (Model 3A), de toevoeging van de interactie afnamejaar*boeken (Model 4A) en een random slope de interactie tussen aantal boeken thuis en afnamejaar (Model 4B): veranderingen tussen 2003 en 2011.

	Model 3A		Model 3B		Model 4A		Model 4B	
	b	S.E.	b	S.E.	b	S.E.	b	S.E.
Intercept	536.91***	6.17	537.34***	6.04	536.76***	6.26	536.07***	6.16
Afnamejaar	3.37	1.93	2.91	1.96	4.20	2.97	4.31	2.96
<i>Leerlingniveau</i>								
Geslacht	6.79***	1.20	6.79***	1.19	6.79***	1.19	6.80***	1.20
Geboortemaand	-.48*	.20	-.49*	0.20	-.49*	.20	-.49*	.20
Normaaltvorderend	36.03***	1.55	36.05***	1.53	36.05***	1.53	36.05***	1.54
Aantal boeken thuis	7.27***	.80	7.17***	0.70	7.45***	.82	7.47***	.81
Thuistaal	-8.88***	1.41	-8.85***	1.38	-8.85***	1.38	-8.86***	1.38
<i>Interactietermen</i>								
Afnamejaar*boeken					-.59	1.02	-.63	1.02
<i>Schoolniveau</i>								
Compositie SES	6.72*	2.89	6.56*	2.79	6.54*	2.79	6.85*	2.72
Compositie thuistaal	-38.52***	4.91	-36.54***	4.89	-36.58***	4.89	-36.51***	4.92
<i>Variantie (random gedeelte)</i>								
Leerling (intercept)	2561.46	54.49	2535.31	55.37	2535.36	55.38	2535.05	55.22
Klas (intercept)	143.12	35.39	144.60	35.60	144.70	35.58	143.53	34.87
<i>School</i>								
Intercept	174.42	39.76	356.32	99.17	355.02	99.37	333.84	98.09
Slope aantal boeken thuis			25.46	13.30	25.31	13.37	29.78	18.60
Slope afnamejaar							104.47	174.06
Slope afnamejaar*boeken							3.54	9.24

* p < .05; ** p < .01; *** p < .001

Ditzelfde gaan we na voor de verschillen tussen scholen betreffende de samenhang tussen thuistaal en wiskundeprestaties. Model 3B in Tabel 17 toont het model waarbij de thuistaal als random slope wordt opgenomen. We zien dat dit model een beter model vormt dan Model 3A (Tabel 16). Dit betekent dat, wanneer de data van 2003 en 2011 tesamen bekeken wordt, er significante verschillen bestaan tussen scholen in de samenhang tussen thuistaal en wiskundeprestaties. In Model 4A voegen we vervolgens opnieuw de interactieterm 'thuistaal * afnamejaar' toe. Dit model is geen beter model dan Model 3B, wat aangeeft dat de samenhang tussen thuistaal en wiskundeprestaties niet verschillend is in beide afnamejaren. Tot slot voegen we een random slope van de interactieterm toe aan het model (Model 4B), m.a.w. laten we toe dat de verschillen tussen scholen in de samenhang tussen thuistaal en afnamejaar verschillend is voor 2003 en 2011. Tabel 16 toont aan dat ook hier geen beter model bekomen wordt dan Model 4A, wat betekent dat de verschillen tussen scholen in de samenhang van thuistaal en wiskundeprestaties niet gewijzigd zijn in 2011 ten opzichte van 2003.

Tabel 16. Model fit van random slopes op schoolniveau, voor thuistaal en interactieterm 'thuistaal * afnamejaar'.

	2003-2011			
	Log Likelihood	Chi-square	Df	p
Model 3A	-50255.6			
Model 3B: Random slope thuistaal	-50247.7	15.8	2	.000
Model 4A: toevoeging interactieterm 'thuistaal*afnamejaar'	-50246.3	.2	1	.655
Model 4B: Random slope interactieterm	-50244.6	3.46	7	.839

* p < .05; ** p < .01; *** p < .001

Tabel 17. Invloed aantal boeken thuis, thuistaal en schoolcompositie op wiskundeprestaties (model 3A), met een random slope voor thuistaal (Model 3B), toevoeging van de interactieterm afnamejaar*thuistaal (Model 4A) en een model met random slope voor de interactie tussen thuistaal en afnamejaar (Model 4B): veranderingen tussen 2003 en 2011.

	Model 3A		Model 3B		Model 4A		Model 4B	
	b	S.E.	b	S.E.	b	S.E.	b	S.E.
Intercept	536.91 ^{***}	6.17	539.81 ^{***}	6.51	540.19 ^{***}	6.45	540.32 ^{***}	6.36
Afnamejaar	3.37	1.93	3.57	1.95	2.84	2.08	2.89	2.10
<i>Leerlingniveau</i>								
Geslacht	6.79 ^{***}	1.20	6.74 ^{***}	1.19	6.78 ^{***}	1.18	6.76 ^{***}	1.19
Geboortemaand	-.48 [*]	.20	-.51 [*]	.29	-.51 [*]	.20	-.51 [*]	.20
Normaalvorderend	36.03 ^{***}	1.55	35.78 ^{***}	1.52	35.76 ^{***}	1.52	35.72 ^{***}	1.51
Aantal boeken thuis	7.27 ^{***}	.80	7.22 ^{***}	.81	7.21 ^{***}	.81	7.22 ^{***}	.81
Thuistaal	-8.88 ^{***}	1.41	-9.99 ^{***}	1.28	-12.33 ^{***}	1.77	-12.88 ^{***}	1.71
<i>Interactietermen</i>								
Afnamejaar*thuistaal					4.55	2.62	5.19 [*]	2.53
<i>Schoolniveau</i>								
Compositie SES	6.72 [*]	2.89	5.57	2.97	5.56	2.95	5.51	2.89
Compositie thuistaal	-38.52 ^{***}	4.91	-40.84 ^{***}	5.57	-41.25 ^{***}	5.52	-41.52 ^{***}	5.40
<i>Variantie random gedeelte</i>								
Leerling (intercept)	2561.46	54.49	2541.40	52.83	2540.22	52.77	2541.50	52.84
Klas (intercept)	143.12	35.39	143.38	35.15	143.19	35.22	140.82	33.93
School								
Intercept	174.42	39.76	162.71	37.83	162.41	37.74	147.00	40.73
Slope thuistaal			100.84	51.90	105.22	50.95	157.00	64.31
Slope afnamejaar							58.12	116.12
Slope afnamejaar*thuistaal							52.33	128.72

*p < .05; **p < .01; ***p < .001

4. Referenties

- Belfi, B., Arkens, T., Van Damme, J., & Gielen, S. (2012). *Wiskunde in het lager onderwijs in TIMSS 2003 en TIMSS 2011 in Vlaanderen, zonder internationaal perspectief: Een voorlopige vergelijking*. Leuven: KU Leuven, Centrum voor Onderwijseffectiviteit en –evaluatie.
- Gielen, S., Bellens, K., Belfi, B., & Van Damme, J. (2012). *Het vierde leerjaar basisonderwijs in Vlaanderen: Resultaten van TIMSS 2011 in international perspectief en in vergelijking met TIMSS 2003*. Leuven: KU Leuven, Centrum voor Onderwijseffectiviteit en –evaluatie.
- Meelissen, M.R.M., & Drent, M. (2008). *TIMSS-2007 Nederland. Trend in leerprestaties in exacte vakken in het basisonderwijs*. Enschede: Universiteit Twente, Vakgroep Onderwijsorganisatie en –management.
- Ministerie van Onderwijs en Vorming (2012). *Gelijke onderwijskansen*. Opgevraagd op 26 Februari 2013 via <http://www.ond.vlaanderen.be/gok/>
- Van Damme, D. (2012). *Excellent onderwijs voor iedereen in de 21^{ste} eeuw*. Toespraak gehouden op de VLOR-startdag, Brussel, 27 september 2012.
- Van Damme, J., Van Den Broeck, A., Brusselmans-Dehairs, C., & Valcke, M. (2004). *Vlaanderen in TIMSS 2003*. Brussel: Ministerie van de Vlaamse Gemeenschap.
- von Davier, M., Gonzalez, E., & Mislevy, R.J. (2009). What are plausible values and why are they useful? In IEA-ETS, *IERI Monograph Series: Issues and methodologies in large-scale assessments* (Vol. 2, pp. 9-36). US: IEA-ETS Research Institute.
- TIMSS (2011). *TIMSS 2011 item parameters from concurrent calibration*. Boston: IEA TIMSS & PIRLS International Study Center.

Bijlagen

Bijlage 1. Steekproeftrekking, uiteindelijke steekproef en uitsluitingen in Vlaanderen in TIMSS 2011

In deze bijlage geven we uitleg bij de steekproeftrekking in TIMSS 2011 en de uiteindelijke steekproef. Opmerkelijk hierbij is dat er in vergelijking met TIMSS 2003 veel leerlingen uitgesloten worden uit de uiteindelijke, internationale database. Daarom gaan we in een derde paragraaf verder in op deze substantiële groep van uitgesloten leerlingen in TIMSS 2011.

1. Steekproeftrekking¹¹

Uit de Vlaamse basisscholen met een vierde leerjaar werd een steekproef van 150 scholen getrokken, waarbij representativiteit voor het officieel versus vrij onderwijs en voor de provincies werd nagestreefd. Franstalige basisscholen in Vlaanderen en scholen met minder dan 5 leerlingen in het vierde leerjaar werden vooraf uitgesloten uit de onderzoekspopulatie en konden dus niet in de steekproef terecht komen. In Vlaanderen ging het om 9 Franstalige scholen en 12 scholen met minder dan vijf leerlingen. Het gaat hier om een totaal van 350 leerlingen (0.55%) die werden uitgesloten.

Zoals gezegd werd bij de steekproeftrekking rekening gehouden met de provincie waarin de scholen gelegen zijn (Antwerpen, Brussel en Vlaams-Brabant, Limburg, Oost-Vlaanderen en West-Vlaanderen), alsook met het net waartoe ze behoren (officieel of vrij onderwijs). Bijkomend werd ook het aandeel lage SES-leerlingen bekeken. Door met deze drie criteria (provincie, net en aandeel lage SES-leerlingen) rekening te houden, werd ervoor gezorgd dat de steekproef op die punten representatief is voor de Vlaamse scholen en leerlingen.

¹¹ Gebaseerd op Gielen, Bellens, Belfi & Van Damme (2012)

2. Uiteindelijke steekproef

Van de 150 aanvankelijk aangesproken scholen bleken er 114 bereid tot medewerking. Daarnaast werden nog 28 reservescholen bereid gevonden deel te nemen. Dus beschikken we over resultaten voor 142 scholen. Voor het eerst werden daarnaast ook 6 BLO-scholen (met type 3 en/of 8) gevraagd om mee te doen. Vijf van die scholen hebben inderdaad ook meegewerkt. Omdat het om een eerste proef ging en vooral omdat de BLO-leerlingen niet in het vierde leerjaar zitten, werden de gegevens uiteindelijk niet betrokken in de internationale vergelijking. Omdat BLO-scholen niet echt een vierde leerjaar hebben, besloot de I.E.A. in 2011 immers om de BLO-scholen niet tot de populatie te rekenen.

Aan leerkrachten van de deelnemende klassen werd steeds gevraagd om een klassenlijst in te vullen en hierbij aan te duiden of een leerling volgens hen officieel uitgesloten moest worden van deelname aan de toets omwille van (1) een fysieke handicap, (2) ernstige leer- en/of leesproblemen, of (3) een onvoldoende beheersing van de Nederlandse taal. Leerlingen van wie de leerkracht dit aangeeft, worden in de internationale vergelijking niet mee opgenomen. Zoals blijkt uit Tabel 18 werden in totaal 197 leerlingen of 4.49% van de getrokken steekproef uitgesloten uit de internationale vergelijkingen. Vermeldenswaard is dat in 2011 significant meer jongens dan meisjes uitgesloten werden op basis van ernstige leer- en/of leesproblemen of omdat zij de Nederlandse taal onvoldoende beheersen. Belangrijk om vermelden is ook dat in 2003 veel minder individuele leerlingen uitgesloten werden (nl. 0.4%).

Tabel 18. Uitsluitingen individuele leerlingen in TIMSS 2011, opgesplitst naar geslacht en reden van uitsluiting.

	Totaal	Meisjes	Jongens
Fysieke handicap	2	1	1
Ernstige leer- en/of leesproblemen	126	53	73
Beheersing Nederlandse taal onvoldoende	69	27	42
Totaal	197	81	116

In totaal bestaat de steekproef in TIMSS 2011 hierdoor uit 4849 leerlingen, verdeeld over 286 klassen in 142 scholen. In 2003 ging het over 4712 leerlingen, verdeeld over 258 klassen in 149 scholen.

3. Uitsluitingen

Omdat de focus van dit rapport ligt op een vergelijking van de Vlaamse resultaten in 2011 met die van 2003, gaan we nog iets grondiger in op het grote verschil tussen beide meetmomenten inzake het percentage leerlingen dat van het onderzoek uitgesloten werd. We beperken ons hier tot de twee grootste groepen.

In beide jaren werd het BLO buiten beschouwing gelaten. In 2003 impliceerde dit het buiten beschouwing laten van 6,1 % van de leeftijdsgroep van 10 jaar, nl. 4,6 % van de meisjes en 7,6 % van de jongens; in 2011 impliceerde dit het uitsluiten van 6,9 % van de leerlingen, nl. 5,3 % van de meisjes en 8,4 % van de jongens¹². Door de leerkrachten van de deelnemende scholen werden in 2003 0,4 % van de individuele leerlingen uitgesloten, zowel in totaal als per geslacht. In 2011 was dit percentage zoals gezegd heel wat hoger, nl. 4 % van de meisjes en 5 % van de jongens.

Als we gemakkelijksshalve de percentages BLO-leerlingen en de individuele uitsluitingen optellen, komen we voor 2003 op 4,6 plus 0,4 of 5 % van de meisjes en op 7,6 plus 0,4 of 8 % van de jongens. In 2011 ging het om 5,3 plus 4 of 9,3 % van de meisjes (d.i. een toename van 4,3 %) en om 8,4 plus 5 of 13,4 % van de jongens (d.i. een toename van 5,4 %). In de totale groep bedraagt de toename 4,9 % (6,5 versus 11,4). Die zeer substantiële groep uitgesloten Vlaamse leerlingen zal extra aandacht vragen bij eventuele toekomstige participaties aan zo'n internationaal onderzoek.

Leerlingen van wie de leerkracht aangeeft dat zij uitgesloten dienen te worden, vullen soms wel de prestatietoetsen en/of de achtergrondvragenlijst in, maar worden verwijderd uit de finale databank. Hierdoor worden ze in de internationale vergelijking niet mee opgenomen maar hebben we van een aantal onder hen wel prestatiegegevens. Door middel van deze gegevens gaan we in deze paragraaf dieper in op de resultaten van de leerlingen uit BLO, die als groep werden uitgesloten en de resultaten van de leerlingen die op individuele basis werden uitgesloten.

We hebben prestatiegegevens van wiskunde van alle BLO-leerlingen en van 151 (van de 197) individueel uitgesloten leerlingen. Dit betekent dat we voor 46 leerlingen die individueel werden uitgesloten, geen prestatiegegevens hebben. Van de BLO-leerlingen en de 151 individueel uitgesloten leerlingen hebben we echter enkel ruwe prestatiegegevens (met name de antwoorden op de toetsen) voorhanden en werden geen plausible values (zie ook Bijlage 4) berekend. Daarom dienden we zelf prestatiescores te berekenen, zodat we ook voor de BLO-leerlingen en de individueel uitgesloten leerlingen een prestatie-indicator ter beschikking hebben. Op basis van de geschatte

¹² Percentages beschikbaar gesteld door G. Van Landeghem.

item parameters (zie TIMSS, 2011) is het mogelijk om een IRT-score te berekenen voor alle leerlingen waarvan we toetsscores beschikbaar hebben¹³. In Tabel 19 geven we het gemiddelde prestatieniveau weer van de leerlingen die wel werden opgenomen in de internationale databank, gebaseerd op de zelf berekende IRT-score versus op de plausible values. Opvallend is dat de plausible values de vaardigheid wat hoger inschatten. Vermoedelijk komt dit doordat bij de berekening van de plausible values rekening gehouden wordt met achtergrondkenmerken van leerlingen (zie Bijlage 4), wat niet in rekening wordt gebracht in de aanvankelijke berekening van de IRT-scores.

Tabel 19. Gemiddelde prestatieniveau van leerlingen opgenomen in internationale databank, op basis van de zelf berekende IRT-scores en de plausible values (PV) berekend door TIMSS.

	IRT-score	PV1	PV2	PV3	PV4	PV5
Gemiddelde	545.56	550.78	551.36	550.27	549.88	550.66
S.D.	63.13	59.59	59.36	58.36	59.34	59.29

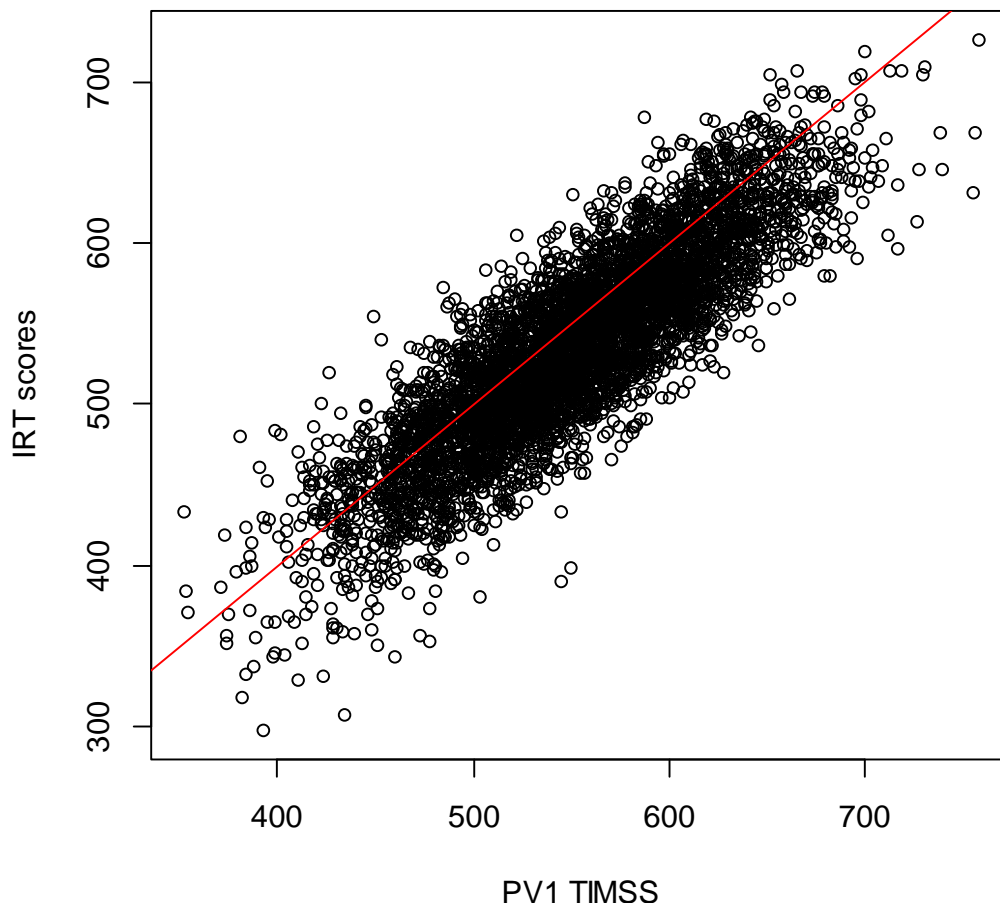
Hoewel deze zelf berekende IRT-scores niet volledig overeenstemmen met de plausible values die berekend worden door TIMSS, zien we in Grafiek 21 dat de IRT-scores van de leerlingen die opgenomen werden in de internationale databank hoog correleren (met name $R = .86$) met de voor hen berekende plausible value van TIMSS. Omdat geen beter alternatief voorhanden is als indicator van de prestaties van BLO-leerlingen en individueel uitgesloten leerlingen, maken we daarom gebruik van de IRT-scores om de prestaties van deze groepen leerlingen te vergelijken met de prestaties van de leerlingen die werden opgenomen in de internationale databank.

Door middel van deze IRT-scores kunnen we de prestatiescores van de BLO-leerlingen en van de individueel uitgesloten leerlingen dus vergelijken met de prestatiescores van de leerlingen die opgenomen werden in de internationale databank. De resultaten staan weergegeven in Tabel 20 en Grafiek 22. De resultaten van de leerlingen die niet in de databank werden opgenomen wijken significant af van de prestaties van de leerlingen die opgenomen werden in de internationale databank: zowel de BLO-leerlingen als de individueel uitgesloten leerlingen behalen significant lagere

¹³ Deze IRT-scores wijken (lichtjes) af van de IRT-scores die gebruikt werden in het rapport van Belfi et al. (2012). De IRT-scores in Belfi et al. (2012) werden berekend op basis van de Vlaamse data in 2003 en 2011. Voor de berekening van de plausible values werd een kalibratie uitgevoerd met de data van alle landen samen over twee opeenvolgende data afnamejaren (voor de meeste landen is dit 2007-2011). De nu berekende IRT-scores maken gebruik van de kalibratieparameters die TIMSS heeft gebruikt (deze waren nog niet beschikbaar bij het berekenen van de IRT-scores voor het rapport van Belfi et al., 2012). Hierdoor kunnen deze scores licht van elkaar afwijken. Dit betekent dat de huidige IRT-scores sterker overeen komen met de plausible values die berekend werden in TIMSS 2011.

prestaties dan de leerlingen die opgenomen werden in de internationale databank. Opmerkelijk is dat de individueel uitgesloten leerlingen op hun beurt significant lagere prestaties behalen dan de BLO-leerlingen. Dit is voornamelijk te verklaren door de hogere leeftijd van de BLO-groep.

Grafiek 21. Correlatie tussen de eerste plausible value berekend door TIMSS en de nieuw berekende IRT-scores van de leerlingen die opgenomen werden in de internationale databank.



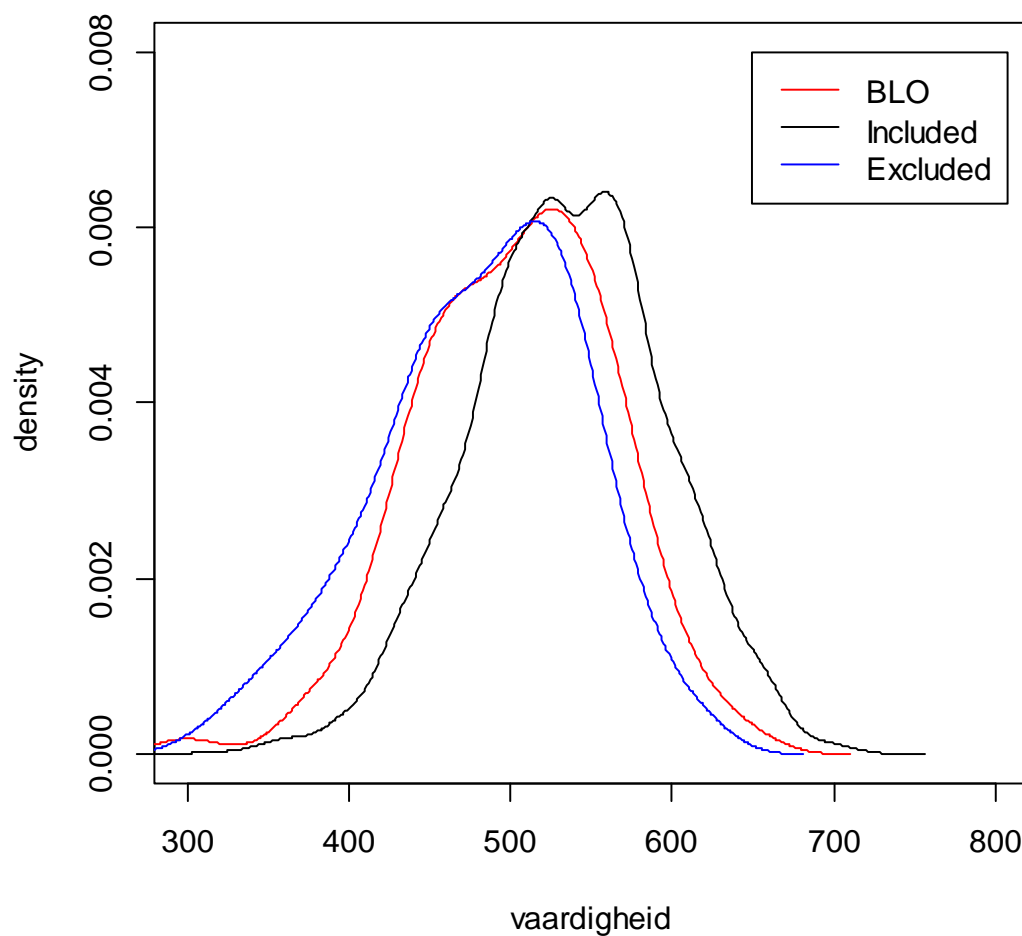
Zoals in paragraaf 2 van deze bijlage werd aangegeven, werden BLO-leerlingen momenteel niet beschouwd als leerlingen die in het vierde leerjaar zitten. Dit leidt echter tot een hoog percentage leerlingen die uitgesloten worden in Vlaanderen, wat ten koste gaat van de representativiteit van de verzamelde gegevens. Daarom bekijken we ook de gemiddelde prestaties in Vlaanderen indien de BLO-leerlingen wel zouden opgenomen worden in de internationale databank. Stel dat we de BLO-leerlingen mee zouden opnemen in de internationale vergelijking, dan zou het gemiddelde prestatieniveau 543.85 in plaats van 545.56 zijn (op basis van de zelf berekende IRT-scores, die zoals gezegd negatief afwijken van de internationale prestatiescore). Dit zou geen significante daling

betekenen van het Vlaams gemiddelde met betrekking tot wiskundeprestaties. Vermeldenswaard is de zeer grote overlap in de resultaten van de drie groepen, die blijkt uit Grafiek 22.

Tabel 20. Vergelijking prestaties BLO-leerlingen en individueel uitgesloten leerlingen met prestaties leerlingen opgenomen in internationale databank, TIMSS 2011.

	N	Gemiddelde (S.D.)
Leerlingen in internationale databank	4846	545.56 (63.13)
Individueel uitgesloten leerlingen	151	489.17 (64.26)
BLO-leerlingen	107	513.15 (63.02)

Grafiek 22. Verdeling van de prestatiescores van BLO-leerlingen, individueel uitgesloten leerlingen en leerlingen opgenomen in internationale databank.



Bijlage 2. Wiskundig zelfconcept: een vergelijking tussen 2003 en 2011¹⁴

In de TIMSS leerlingvragenlijst van 2003 en 2011 werd door middel van vier items getoetst hoe zelfzeker de leerlingen waren over hun eigen wiskundige vaardigheden. Deze items zijn: (1) Meestal ben ik goed in wiskunde, (2) Wiskunde is voor mij moeilijker dan voor veel van mijn klasgenoten^(*), (3) Ik ben gewoon niet goed in wiskunde^(*) en (4) Ik leer vrij snel dingen over wiskunde. Er waren telkens 4 antwoordcategorieën, gaande van 'helemaal niet akkoord' tot 'helemaal akkoord'¹⁵. Het gemiddelde van de vier items werd genomen als indicatie van het wiskundig zelfconcept van de leerling ($\alpha_{2003} = 0.82$; $\alpha_{2011} = 0.82$). Hoe hoger de score op deze schaal, hoe positiever een leerling over zijn/haar wiskundige vaardigheden denkt.

1. Vergelijking tussen het gemiddelde wiskundig zelfconcept in 2003 en 2011

In Tabel 21 wordt het gemiddelde op de schaal 'wiskundig zelfconcept' van TIMSS 2003 en TIMSS 2011 weergegeven. De gemiddelde scores verschillen significant van elkaar. Gemiddeld genomen denken de leerlingen in 2011 positiever over hun eigen wiskundevaardigheden dan in 2003.

Tabel 21. Gemiddeld wiskundig zelfconcept in 2003 en 2011.

	2003		2011		t-toets
	Gemiddelde	SD	Gemiddelde	SD	2003 ≠ 2011
Wiskundig zelfconcept	2.99	.67	3.10	.75	-7.49*

* p < .05; ** p < .01; *** p < .001

¹⁴ Gebaseerd op Belfi et al. (2012)

¹⁵ De negatieve stellingen, aangegeven met een ^(*), werden voor de creatie van de schaal omgezet naar positieve stellingen.

2. Vergelijking tussen het gemiddelde wiskundig zelfconcept van jongens en meisjes in 2003 en 2011

In Tabel 22 wordt het gemiddelde wiskundig zelfconcept weergegeven van jongens en meisjes in 2003 en 2011. In beide afnamejaren hebben jongens een significant hoger wiskundig zelfconcept dan meisjes. Deze bevinding is overeenkomstig met wat doorgaans in onderzoek wordt vastgesteld. Vermeldenswaard is dat het verschil een toenemende tendens vertoont. Daarnaast ligt zowel het wiskundig zelfconcept van de jongens als dat van de meisjes significant hoger in 2011 dan in 2003.

Tabel 22. Gemiddeld wiskundig zelfconcept van jongens en meisjes in 2003 en 2011, getest met een ongepaarde t-toets of het verschil tussen jongens en meisjes significant is in 2003 en 2011, en of het verschil overheen de jaren significant is voor jongens en meisjes.

	2003		2011		t-toets
	Gemiddelde	SD	Gemiddelde	SD	2003 ≠ 2011
Jongens	3.12	.65	3.26	.32	-6.77**
Meisjes	2.87	.68	2.94	.77	-3.62**
t-toets jongens ≠ meisjes	13.14 **		14.98 **		

* p < .05; ** p < .01; *** p < .001

3. Vergelijking van het gemiddelde wiskundig zelfconcept van leerlingen die al dan niet thuis Nederlands spreken tussen 2003 en 2011

In Tabel 23 wordt het gemiddelde wiskundig zelfconcept weergegeven van leerlingen die thuis altijd (of bijna altijd), soms en nooit Nederlands spreken in 2003 en 2011. Als eerste valt op dat alle leerlingen die thuis (bijna) altijd of soms Nederlands spreken een significant hoger wiskundig zelfconcept vertonen in 2011 dan in 2003. Het wiskundig zelfconcept van leerlingen die thuis nooit Nederlands spreken verschilt echter niet tussen 2003 en 2011, mede omwille van de kleine aantallen.

Wanneer de drie categorieën onderling worden vergeleken, blijkt dat in beide afnamejaren leerlingen die thuis nooit Nederlands spreken een nagenoeg even hoog wiskundig zelfconcept hebben als leerlingen die thuis (bijna) altijd Nederlands spreken. Leerlingen die thuis slechts soms Nederlands spreken vertonen in beide afnamejaren een significant lager wiskunde zelfbeeld dan de leerlingen die thuis altijd Nederlands spreken. Het wiskundig zelfconcept van leerlingen die soms en

nooit Nederlands spreken verschilt overigens niet significant, eveneens vooral omwille van de kleine aantallen.

Tabel 23. Gemiddeld wiskundig zelfconcept van leerlingen die thuis altijd (of bijna altijd), soms en nooit thuis Nederlands spreken in 2003 en 2011, getest met een ongepaarde t-toets of het verschil tussen de verschillende categorieën significant is in 2003 en 2011, en of het verschil overheen de jaren significant is per categorie.

	2003		2011		t-toets
	Gemiddelde	SD	Gemiddelde	SD	2003 ≠ 2011
(Bijna) altijd	3.00	.66	3.11	.76	-6.72**
Soms	2.92	.69	3.04	.72	-3.00**
Nooit	3.02	.72	3.10	.77	-1.02
t-toets					
Altijd-Soms	2.16*		2.54*		
Soms-Nooit	-1.28		-.97		
Altijd-Nooit	-.19		.16		

* p < .05; ** p < .01; *** p < .001

4. Vergelijking tussen het gemiddelde wiskundig zelfconcept van leerlingen met verschillende hoeveelheden boeken thuis in 2003 en 2011

In Tabel 24 wordt het gemiddelde wiskundig zelfconcept weergegeven van leerlingen die thuis geen of bijna geen boeken hebben (0-10 boeken), genoeg boeken hebben om een boekenplank mee te vullen (11-25 boeken), genoeg boeken hebben om een boekenkast mee te vullen (26-100 boeken), genoeg boeken hebben om twee boekenkasten mee te vullen (101-200 boeken) en genoeg boeken hebben om drie of meer boekenkasten mee te vullen (>200 boeken) in 2003-2011.

Als eerste valt op dat op alle leerlingcategorieën, behalve de leerlingen die thuis de minste boeken (0-10 boeken) en de meeste boeken (>200 boeken) hebben, over een significant hoger wiskundig zelfconcept beschikken in 2011 dan in 2003. Wanneer we kijken naar de categorieën onderling, valt op dat in zowel 2003 als 2011 de leerlingen met de minste boeken thuis (0-10 boeken) gemiddeld het laagste wiskundig zelfconcept hebben en de leerlingen met de meeste boeken thuis (≥ 200 boeken) gemiddeld het hoogste wiskundig zelfconcept.

In allebei de afnamejaren blijken verder de laagste twee groepen nauwelijks van elkaar te verschillen in wiskundig zelfconcept. De middelste groep (26-100 boeken) heeft een significant hoger wiskundig

zelfconcept dan de laagste twee groepen in beide afnamejaren. In beide afnamejaren heeft verder de op een na hoogste groep (101-200 boeken) een significant hoger wiskundig zelfconcept dan de middelste groep. Het wiskundig zelfconcept van de hoogste groep (≥ 200 boeken) verschilt niet van dat van de op één na hoogste groep in beide afnamejaren.

Tabel 24. Gemiddeld wiskundig zelfconcept van leerlingen die thuis over verschillende hoeveelheden boeken beschikken in 2003 en 2011, getest met een ongepaarde t-toets of het verschil tussen de verschillende categorieën significant is in 2003 en 2011, en of het verschil overheen de jaren significant is per categorie.

	2003		2011		t-toets
	Gemiddelde	SD	Gemiddelde	SD	2003 \neq 2011
0-10	2.90	.67	3.00	.77	-1.47
11-25	2.93	.66	3.01	.75	-2.63**
26-100	2.99	.66	3.11	.74	-5.56**
101-200	3.06	.69	3.18	.74	-3.31**
>200	3.14	.70	3.23	.76	-1.94
t-toets					
0-10 vs 11-25	-.90		-.83		
11-25 vs 26-100	-2.11*		-3.76**		
26-100 vs 101-200	-2.40*		-2.04*		
101-200 vs >200	-1.92		-1.14		

* $p < .05$; ** $p < .01$; *** $p < .001$

5. Vergelijking tussen het gemiddelde wiskundig zelfconcept van leerlingen met een verschillend geboortjaar in 2003 en 2011

In Tabel 25 wordt het gemiddelde wiskundig zelfconcept weergegeven van leerlingen die één leerjaar vertraagd zijn, twee of meer leerjaren vertraagd zijn, op leeftijd zitten en één of meerdere leerjaren versneld zijn in 2003 en 2011.

Allereerst valt op dat in beide afnamejaren de leerlingen die één leerjaar vertraagd zijn een lager wiskundig zelfconcept vertonen dan leerlingen die 'op leeftijd' zitten. Dit verschil is statistisch significant. Het verschil in wiskundig zelfconcept tussen leerlingen die één en twee of meer leerjaren vertraagd zijn is niet significant verschillend in zowel 2003 als 2011. In beide afnamejaren hebben leerlingen die één of twee leerjaren versneld zijn een hoger wiskundig zelfconcept dan leerlingen die op leeftijd zitten. De voorsprong in wiskundig zelfconcept van versnelde leerlingen is echter veel

groter in 2011 dan in 2003; het verschil tussen ‘versnelde’ leerlingen en leerlingen ‘op leeftijd’ is dan ook alleen significant in 2011. Wanneer we het wiskundig zelfconcept van de vier categorieën leerlingen tussen de afnamejaren met elkaar vergelijken, valt op dat enkel de wiskundig zelfconcepten van de leerlingen die op leeftijd zitten en de versnelde leerlingen significant verschillend zijn tussen 2003 en 2011.

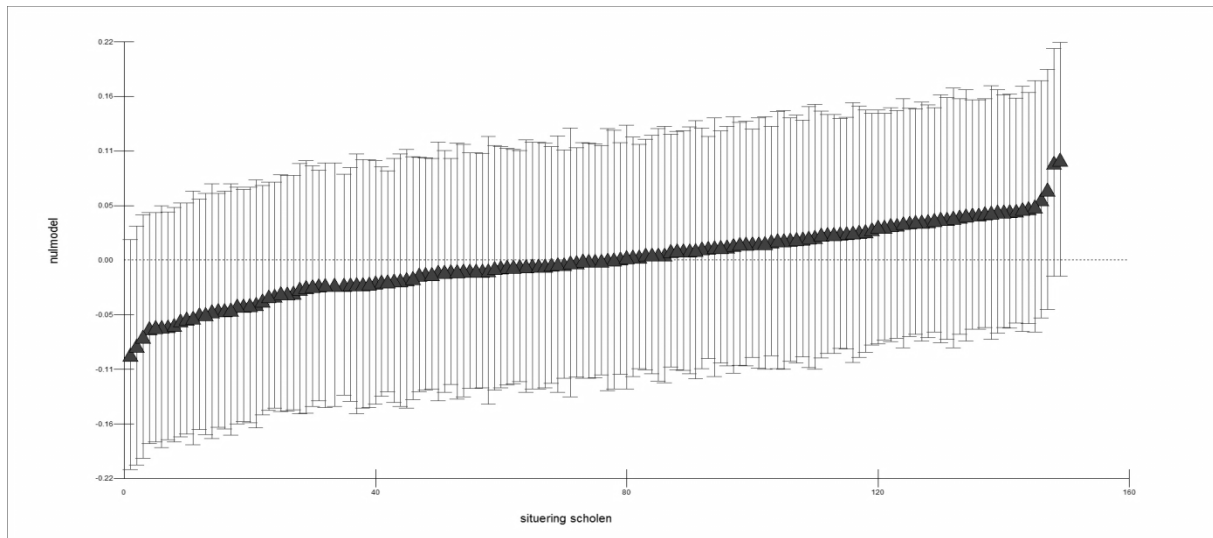
Tabel 25. Gemiddeld wiskundig zelfconcept van leerlingen die 1 of 2 leerjaren vertraagd, versneld, of op leeftijd zijn in 2003 en 2011, getest met een ongepaarde t-toets of het verschil tussen de verschillende categorieën significant is in 2003 en 2011, en of het verschil overheen de jaren significant is per categorie.

	2003		2011		t-toets
	Gemiddelde	SD	Gemiddelde	SD	2003 ≠ 2011
Versneld	3.14	.64	3.51	.61	-3.00**
Normaalvorderend	3.03	.67	3.15	.73	-7.57**
1 jaar vertraging	2.78	.68	2.78	.77	.00
2 of >2 jaar vertraging	2.90	.61	2.93	.72	-.27
t-toets					
Versneld – normaalvorderend	1.03		4.04**		
Normaalvorderend – 1j. vertraging	8.78**		12.05**		
1j. vertraging – 2 of >2j. vertraging	-1.35		-1.80		

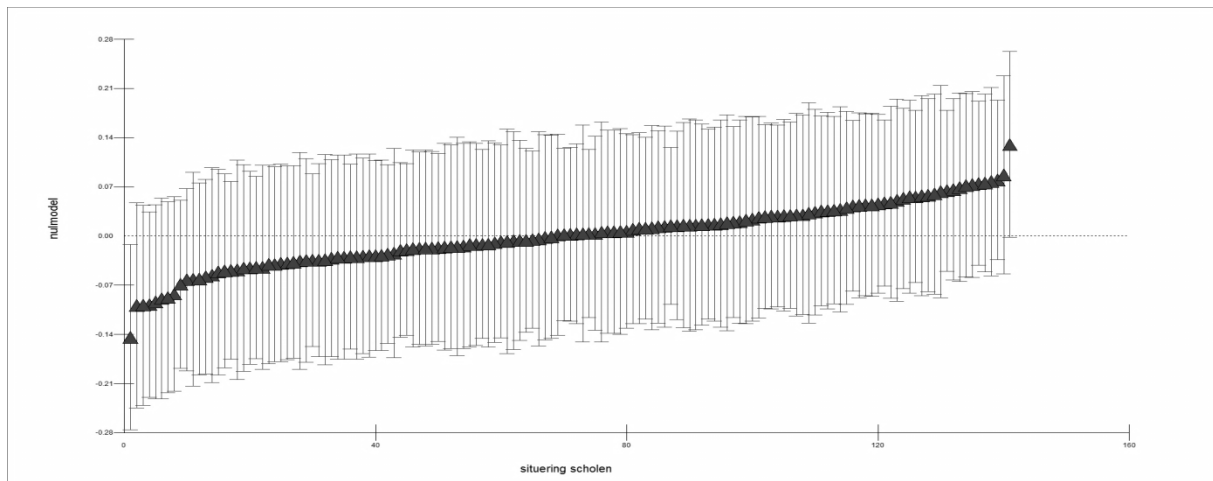
6. Verschillen tussen scholen in wiskundig zelfconcept: Een vergelijking tussen 2003 en 2011

Scholen verschillen niet alleen in gemiddelde wiskunde prestaties, maar ook in gemiddelde wiskundig zelfconcept. Dit heeft opnieuw te maken met de instroomverschillen tussen scholen maar ook met de mate waarin een school aandacht besteedt aan en/of invloed uitoefent op het niet-cognitieve functioneren van leerlingen, waaronder het wiskundig zelfconcept. In Grafiek 23 en Grafiek 24 worden de scholen die aan respectievelijk TIMSS 2003 en TIMSS 2011 hebben deelgenomen, gerangschikt naar hun gemiddelde wiskundig zelfconcept, wanneer er geen correctie wordt toegepast voor de instroomkenmerken.

Grafiek 23. Situering van de deelnemende scholen uit TIMSS 2003 op basis van het wiskundig zelfconcept zonder rekening te houden met de instroomkenmerken van de leerlingen.



Grafiek 24. Situering van de deelnemende scholen uit TIMSS 2011 op basis van het wiskundig zelfconcept zonder rekening te houden met de instroomkenmerken van de leerlingen (exclusief outlier- en BLO-scholen).



Uit Grafiek 23 wordt duidelijk dat er in 2003 geen enkele school afwijkt van het Vlaamse gemiddelde qua wiskundig zelfconcept, wanneer er geen correctie voor instroom-kenmerken plaats vindt. In 2011 is er een school die er uitspringt met een significant lager gemiddeld wiskundig zelfconcept in vergelijking met de andere scholen (Grafiek 24). Opvallend is dat deze school niet significant onderpresteert in wiskunde (wanneer er geen rekening wordt gehouden met instroomverschillen).

Dit suggereert dat een school die effectief is op het gebied van wiskunde, niet noodzakelijk ook effectief hoeft te zijn in het bevorderen van het wiskundig zelfconcept van leerlingen.

In Tabel 26 wordt weergegeven in hoeverre het wiskundig zelfconcept van leerlingen is toe te schrijven aan verschillen tussen leerlingen binnen eenzelfde klas, tussen klassen binnen eenzelfde school en tussen scholen. De resultaten laten zien dat in beide afnamejaren het wiskundig zelfconcept van leerlingen voor ongeveer 99% is toe te schrijven aan verschillen tussen leerlingen binnen eenzelfde klas, helemaal niet is toe te schrijven verschillen tussen klassen binnen eenzelfde school, en voor slechts ongeveer 1% aan verschillen tussen scholen. Het wiskundig zelfconcept van leerlingen is dus iets waar scholen zeer weinig invloed op hebben (of uitoefenen), althans in die zin dat hun gemiddelde invloed op dit punt gelijkaardig is.

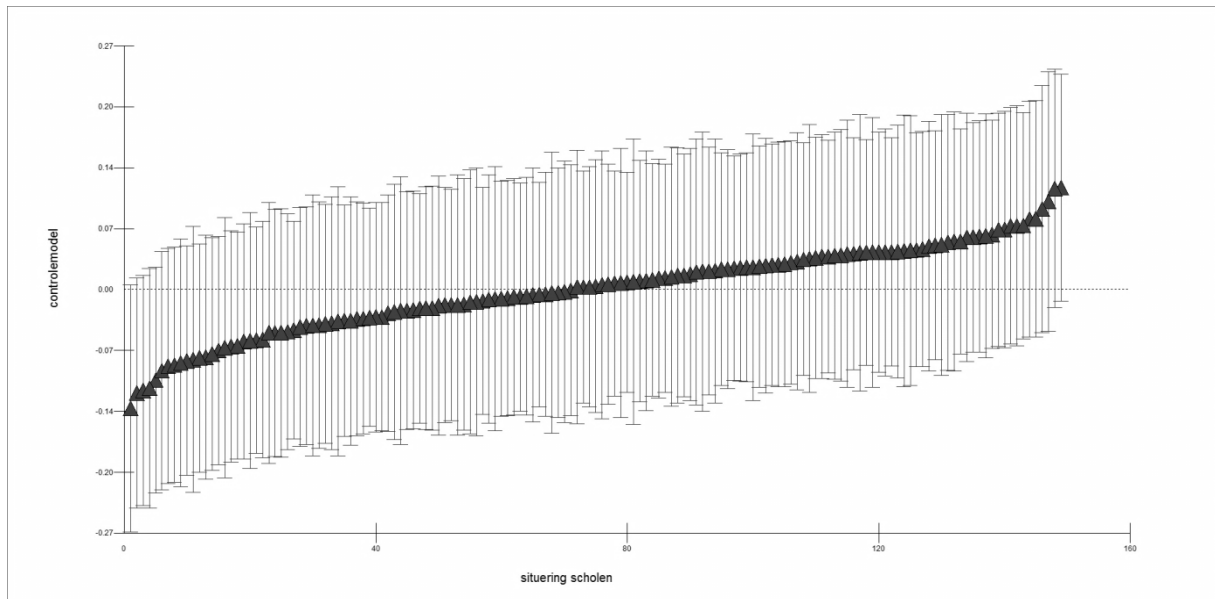
Tabel 26. Multilevel analyse wiskundig zelfconcept: leeg model (Model 0) en model met leerlingvariabelen (Model 1), in 2003 en 2011.

	2003				2011			
	Model 0		Model 1		Model 0		Model 1	
	B	S.E.	B	S.E.	B	S.E.	B	S.E.
Intercept	2.995*	0.012	2.847*	0.045	3.093*	0.013	3.044*	0.038
Sekse			0.262*	0.020			0.318*	0.021
Thuis taal			-0.022	0.022			-0.005	0.021
Aantal boeken thuis			0.049*	0.009			0.054*	0.010
Geboortemaand			-0.011*	0.003			-0.010*	0.003
Geboortjaar			-0.156*	0.023			-0.234*	0.024
Variantie								
School	0.005*	0.002	0.007*	0.002	0.007*	0.003	0.009*	0.003
Klas	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
leerling	0.450*	0.010	0.419*	0.009	0.557*	0.011	0.511*	0.011
Deviantie	9383.4		8870.2		10965.9		10426.5	

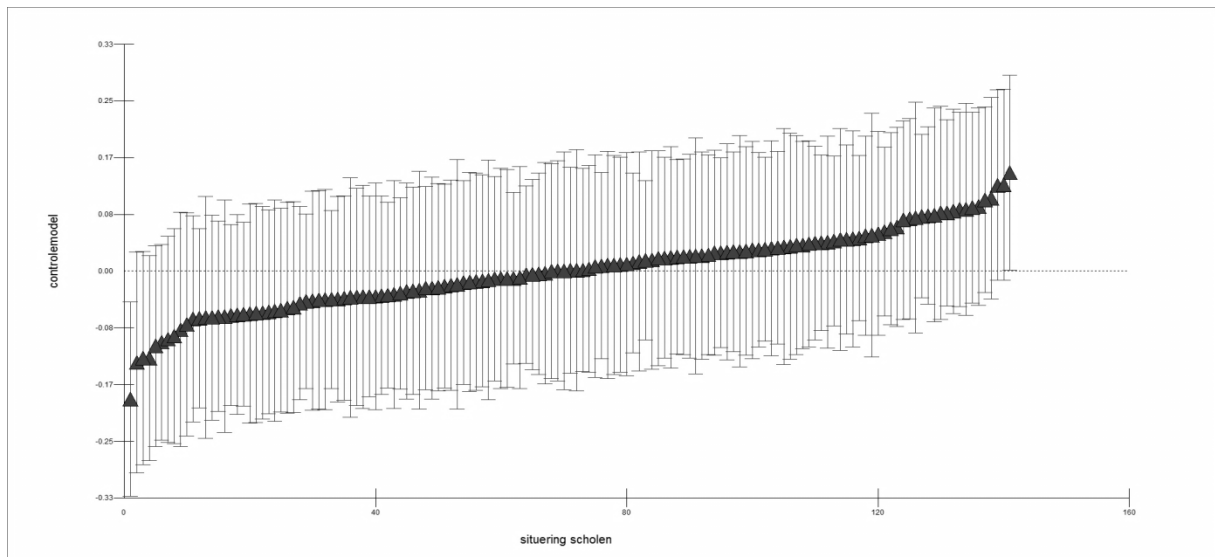
* Deze coëfficiënten zijn statistisch significant bij $\alpha = .05$.

In Grafiek 25 en Grafiek 26 worden de scholen die aan respectievelijk TIMSS 2003 en TIMSS 2011 hebben deelgenomen, gerangschikt naar hun gemiddelde wiskundig zelfconcept, wanneer er correctie wordt toegepast voor de instroomkenmerken. Deze instroomkenmerken zijn: (1) sekse, (2) thuis taal, (3) aantal boeken thuis, (4) geboortemaand en (5) geboortjaar.

Grafiek 25. Situering van de deelnemende scholen uit TIMSS 2003 op basis van het wiskundig zelfconcept, rekening houdend met de instroomkenmerken van de leerlingen.



Grafiek 26. Situering van de deelnemende scholen uit TIMSS 2011 op basis van het wiskundig zelfconcept, rekening houdend met de instroomkenmerken van de leerlingen (exclusief outlier- en BLO-scholen).



Grafiek 26 laat zien dat er ook wanneer er controle plaatsvindt voor de instroomkenmerken, er geen school significant afwijkt van het Vlaamse gemiddelde qua wiskundig zelfconcept in 2003. Grafiek 26 toont verder dat in 2011 na controle voor de instroomkenmerken, de school met het significante ondergemiddelde wiskundig zelfconcept, het nog slechter doet dan wanneer er geen controle plaats

vond voor de instroomkenmerken. Het lagere wiskundig zelfconcept van de leerlingen binnen deze school lijkt dus vooral te wijten te zijn aan processen die zich binnen deze school afspelen. Verder is er nu een school bijgekomen die een significant bovengemiddeld wiskundig zelfconcept bij zijn leerlingen weet te behalen. In Tabel 26, Model 1 worden de coëfficiënten en de standaardfouten van de leerlingkenmerken weergegeven wanneer ze samen in een model zitten. Te zien is dat wanneer alle kenmerken gezamenlijk in een model zitten, alle coëfficiënten, behalve de thuistaal van leerlingen, statistisch significant zijn. Dit indiceert dat de thuistaal van leerlingen geen verband heeft met het wiskundig zelfconcept van leerlingen wanneer er rekening wordt gehouden met de overige leerlingkenmerken.

In de volgende analyses trachten we de variantie in wiskundig zelfconcept op de verschillende niveaus te verklaren door telkens een leerlingkenmerk apart aan het nulmodel toe te voegen. Daarnaast gaan we na hoeveel variantie er door de leerlingkenmerken ‘aantal boeken thuis’ en ‘thuistaal’ gezamenlijk wordt verklaard en hoeveel variantie er door alle leerlingkenmerken tezamen kan worden verklaard. Om een geldige vergelijking tussen de modellen te kunnen maken, hebben we in deze analyses alleen de gegevens gebruikt van leerlingen van wie we informatie hebben over al hun leerlingkenmerken. Zodoende zijn de huidige analyses gebaseerd op de gegevens van 4478 leerlingen in 2003 en 4781 leerlingen in 2011. De resultaten van deze analyses worden weergegeven in Tabel 27 en Tabel 28.

Tabel 27. Percentage verklaarde variantie per leerlingkenmerk in TIMSS 2003.

	<i>Leerling %</i>	<i>Klas %</i>	<i>School %</i>	<i>Totaal %</i>	<i>Deviantie</i>	<i>Verschil in deviantie*</i>
Sekse	3.8%	<i>n.s.</i>	0.0%	3.5%	8988.380	166.426
Thuistaal	0.0%	<i>n.s.</i>	0.0%	0.0%	9152.670	2.136
Aantal boeken thuis	0.9%	<i>n.s.</i>	16.7%	1.1%	9111.363	43.443
Thuistaal + Aantal boeken thuis	0.9%	<i>n.s.</i>	0.0%	0.9%	9110.473	44.333
Geboortemaand	0.4%	<i>n.s.</i>	0.0%	0.4%	9128.960	25.846
Geboortjaar	1.6%	<i>n.s.</i>	0.0%	1.3%	9090.244	64.562
Alle leerlingkenmerken	6.3%	<i>n.s.</i>	0.0%	6.0%	8870.193	284.613

* In vergelijking met het nulmodel

Tabel 28. Percentage verklaarde variantie per leerlingkenmerk in TIMSS 2011.

	<i>Leerling %</i>	<i>Klas %</i>	<i>School %</i>	<i>Totaal %</i>	<i>Deviantie</i>	<i>Verskil in deviantie*</i>
Sekse	4.9%	<i>n.s.</i>	0.0%	4.3%	10593.453	222.177
Thuis taal	0.0%	<i>n.s.</i>	0.0%	0.0%	10811.183	4.447
Aantal boeken thuis	0.9%	<i>n.s.</i>	0.0%	0.9%	10769.955	45.675
Thuis taal + Aantal boeken thuis	0.9%	<i>n.s.</i>	0.0%	0.7%	10767.937	47.693
Geboortemaand	0.4%	<i>n.s.</i>	0.0%	0.4%	10791.661	23.969
Geboortjaar	2.4%	<i>n.s.</i>	14.3%	2.5%	10692.867	122.763
Alle leerlingkenmerken	8.1%	<i>n.s.</i>	0.0%	7.6%	10426.481	389.149

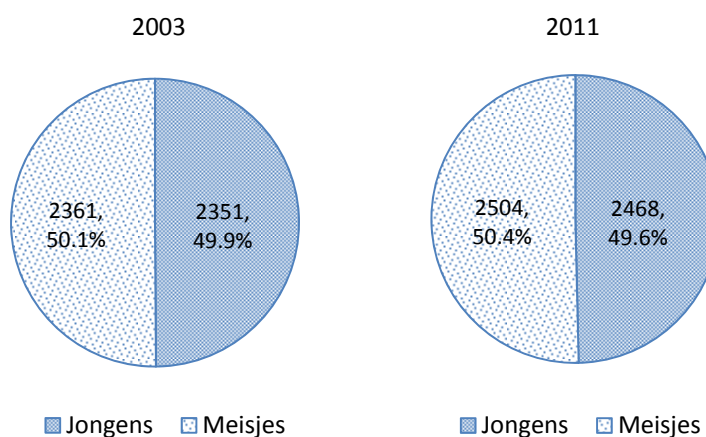
* In vergelijking met het nulmodel.

Tabel 27 en Tabel 28 tonen dat in 2011 alle leerlingkenmerken samen 1.6% meer kunnen verklaren van de totale variantie in het wiskundig zelfconcept van leerlingen. In beide afnamejaren verklaart de sekse van de leerlingen het grootste percentage van de totale variantie in wiskundig zelfconcept en de thuis taal van leerlingen geen. Wat betreft de variantie op leerlingniveau is hetzelfde patroon zichtbaar: de meeste variantie wordt verklaard door de sekse van de leerlingen en geen variantie wordt verklaard door de thuis taal. Op schoolniveau wordt er in beide afnamejaren slechts door één leerlingkenmerk (gedeeltelijk) de variantie in het wiskundig zelfconcept verklaard. In 2003 was dat het 'aantal boeken thuis'. Kennelijk hingen in 2003 verschillen tussen scholen in wiskundig zelfconcept sterker samen met onderlinge verschillen in het gemiddelde aantal boeken dat de leerlingen thuis hebben dan in 2011. In 2011 worden verschillen tussen scholen in wiskundig zelfconcept enkel deels verklaard door onderlinge verschillen in het gemiddelde geboortjaar van leerlingen. Dit duidt wellicht opnieuw op verschillen tussen scholen in percentage zittenblijvers.

Bijlage 3. Verandering inzake leerlingkenmerken¹⁶

1. Geslacht

In Figuur 1 wordt een overzicht gegeven van de aantallen/percentages meisjes en jongens die deelnamen in TIMSS 2003 en TIMSS 2011. Het is duidelijk te zien dat in allebei de afnamejaren het aandeel meisjes/jongens gelijk verdeeld is over de steekproef.

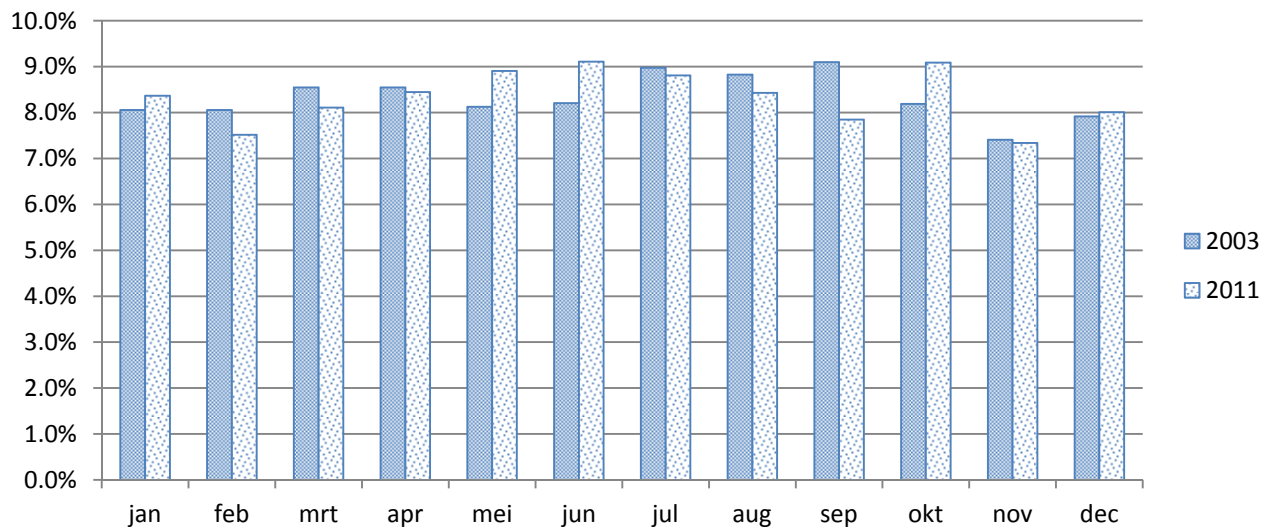


Figuur 1. Verdeling meisjes/jongens in 2003 en in 2011.

¹⁶ Gebaseerd op Belfi et al. (2012)

2. Geboortemaand

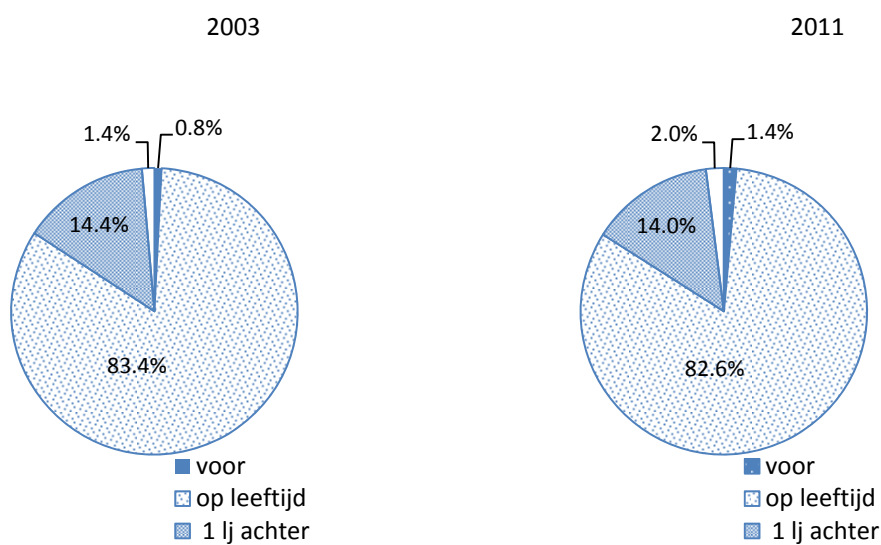
In Figuur 2 wordt een overzicht gegeven van de percentages leerlingen die in de afzonderlijke maanden zijn geboren in 2003 en 2011. Het is duidelijk te zien dat in beide afnamejaren de leerlingen redelijk evenredig verdeeld zijn over de twaalf maanden. Ook tussen de afnamejaren is er sprake van slechts kleine verschillen in percentages per maand.



Figuur 2. Verdeling van leerlingen over de 12 geboortemaanden in 2003 en in 2011.

3. Al dan niet normaalvorderend

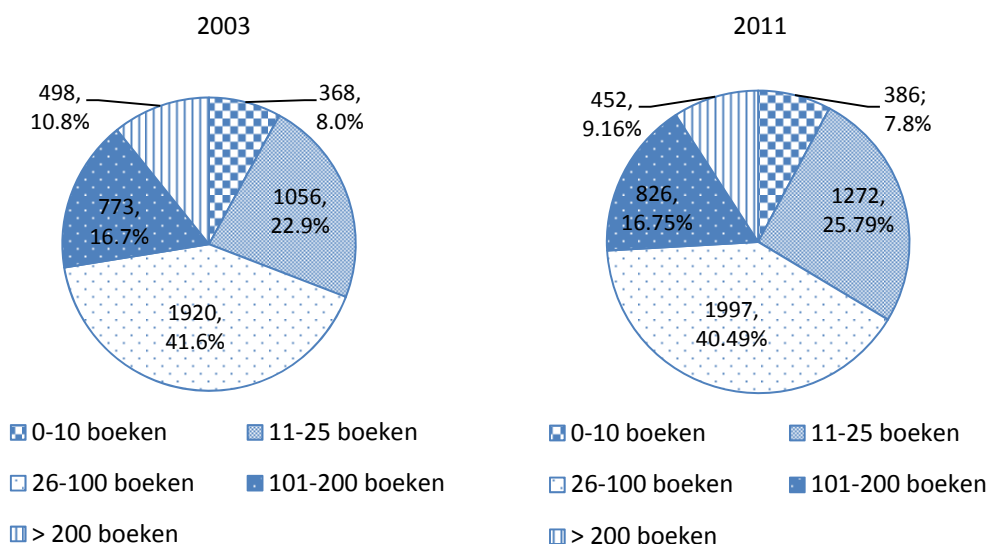
In Figuur 3 wordt getoond welk aandeel leerlingen er één of meerdere leerjaren vertraagd of versneld is de afnamejaren 2003 en 2011. Te zien is dat het merendeel van de leerlingen in beide afnamejaren gewoon 'op leeftijd' zit (+/-83%). In beide afnamejaren bestaat de op één na grootste groep leerlingen uit leerlingen die één leerjaar vertraagd zijn (+/- 14%). De op één na kleinste groep bestaat uit leerlingen die twee of meer leerjaren vertraagd zijn (+/- 2%). Het aandeel leerlingen dat één of meerdere leerjaren versneld is, is in beide afnamejaren zeer klein (+/- 1%). De meest opvallende wijziging is de toename van het aandeel versnelde leerlingen.



Figuur 3. Verdeling van leerlingen die voor (versneld), op leeftijd, 1 leerjaar of 2 leerjaren achter (vertraagd) zijn, in 2003 en 2011.

4. Aantal boeken thuis

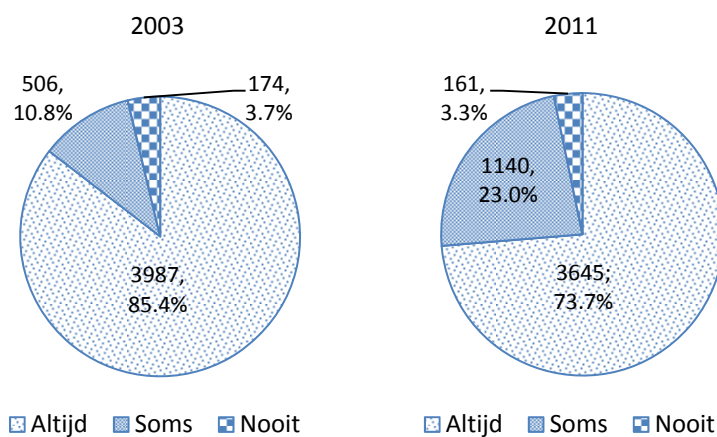
In Figuur 4 wordt een overzicht gegeven van de aantallen/percentages leerlingen die thuis geen of erg weinig boeken hebben (0-10 boeken), genoeg boeken hebben om een boekenplank mee te vullen (11-25 boeken), genoeg boeken hebben om een boekenkast mee te vullen (26-100 boeken), genoeg boeken hebben om twee boekenkasten mee te vullen (101-200 boeken) en genoeg boeken hebben om drie of meer boekenkasten mee te vullen (meer dan 200 boeken). Uit Figuur 4 blijkt allereerst dat de verdeling tussen de 5 categorieën in beide afnamejaren nagenoeg dezelfde is. Het grootste aandeel van de leerlingen beschikt thuis over 26-100 boeken (1 boekenkast), een iets kleiner aandeel van de leerlingen beschikt thuis over 11-25 boeken (1 boekenplank), een nog iets kleiner aandeel van de leerlingen beschikt thuis over 101-200 boeken (2 boekenkasten), hierop volgt een nog kleiner aandeel van leerlingen dat thuis meer dan 200 boeken heeft (>2 boekenkasten). Tenslotte beschikt het kleinste aandeel leerlingen thuis over 0-10 boeken.



Figuur 4. Aantal en percentage leerlingen die thuis geen of zeer weinig boeken hebben (0-10 boeken), genoeg boeken hebben om een boekenplank mee te vullen (11-25 boeken), genoeg boeken hebben om een boekenkast mee te vullen (26-100 boeken), genoeg boeken hebben om twee boekenkasten mee te vullen (101-200 boeken) en genoeg boeken hebben om drie of meer boekenkasten mee te vullen (>200 boeken) in 2003-2011.

5. Thuistaal

In Figuur 5 wordt een overzicht gegeven van de aantallen/percentages leerlingen die thuis altijd (of bijna altijd), soms en nooit thuis Nederlands spreken in 2003 en 2011. Het is duidelijk te zien dat in allebei de afnamejaren het grootste aandeel van de leerlingen thuis (bijna) altijd Nederlands spreekt, een aanzienlijk kleiner aandeel slechts soms thuis Nederlands spreekt en een nog kleiner aandeel thuis nooit Nederlands spreekt. De exacte verdeling tussen de drie groepen verschilt echter per afnamejaar. Hoewel het aandeel leerlingen dat thuis nooit Nederlands spreekt nagenoeg identiek is in beide afnamejaren, is er anno 2011 sprake van een aanzienlijk kleiner aandeel leerlingen dat thuis (bijna) altijd Nederlands spreekt, en een aanzienlijk groter aandeel leerlingen dat slechts soms Nederlands spreekt, dan in 2003.



Figuur 5. Verdeling van leerlingen die thuis (bijna) altijd, soms en nooit Nederlands spreken in 2003-2011.

Bijlage 4. “Plausibele waarden” als indicatoren van prestaties¹⁷

Het doel van grootschalige dataverzamelingen zoals TIMSS, is het verzamelen van data over vaardigheden van leerlingen in verschillende domeinen (o.a. wiskunde, wetenschappen, taal, e.d.) bij een zo representatief mogelijke steekproef van leerlingen. De verzamelde data worden gebruikt om uitspraken te doen over de vaardigheden van bepaalde groepen binnen een populatie (bv: wiskundevaardigheden van jongens, taalvaardigheden van lage SES-leerlingen, wetenschappen-vaardigheden van anderstaligen, e.d.). Om deze uitspraken te kunnen doen, moeten de vaardigheidsscores van de leerlingen in de steekproef die tot een bepaalde groep behoren (bv. lage SES-leerling) een zo goed mogelijke weerspiegeling zijn van de vaardigheidsscores van die gehele groep in de populatie (bv. populatie lage SES-leerlingen). Daarom werkt men in grootschalige dataverzamelingen niet met individuele vaardigheidsscores maar met plausibele waarden.

Plausibele waarden zijn vaardigheidsindicatoren die de verdeling van de vaardigheidsscores van bepaalde groepen van leerlingen in de populatie zo optimaal mogelijk representeren. Plausibele waarden worden berekend op basis van de scores van individuele leerlingen op de toetsen, aangevuld met informatie over achtergrondkenmerken. Elke leerling krijgt op deze manier vijf plausibele waarden toegekend.

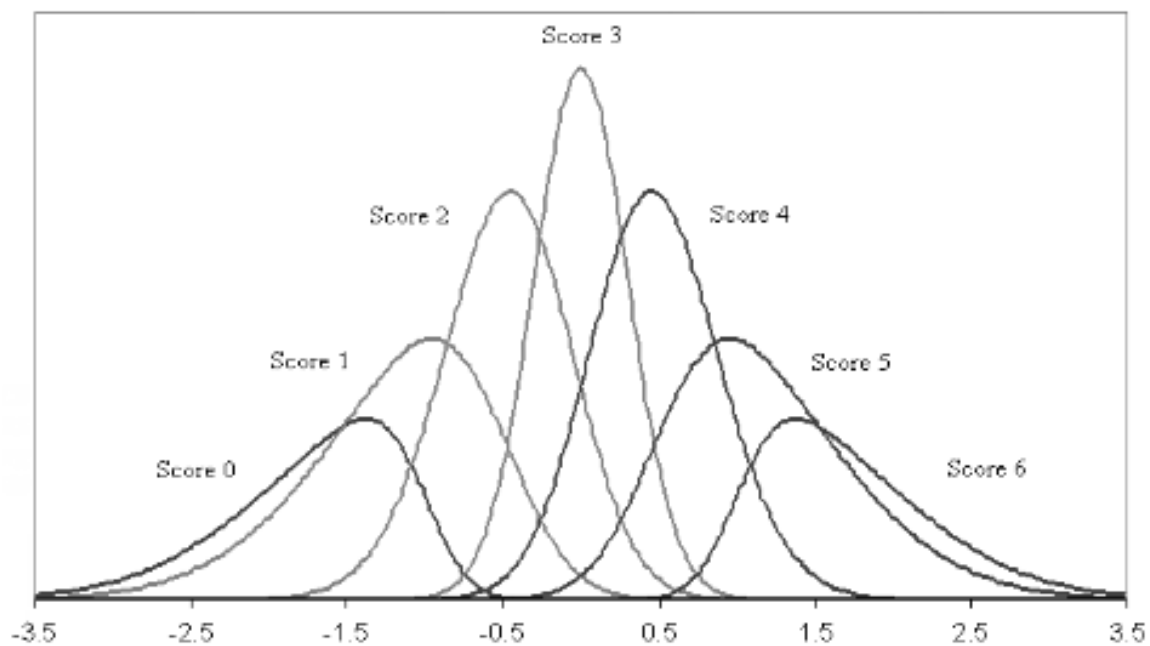
Allereerst worden plausibele waarden berekend op basis van de scores van leerlingen op de items die zij hebben ingevuld. Om uitspraken te kunnen doen over de vaardigheden van leerlingengroepen binnen de populatie, omvatten grootschalige dataverzamelingen een groot aantal items om een bepaalde onderliggende vaardigheid te bepalen. Al deze items samen vormen een valide en betrouwbaar meetinstrument om de onderliggende vaardigheid van een leerling te bepalen. Individuele leerlingen vullen echter niet alle items in. De gehele set van items wordt opgesplitst in blokken of clusters van items en elke leerling vult vervolgens een aantal blokken of clusters in. Dit betekent dat elke leerling slechts een gedeelte van de hele pool van items die peilen naar een bepaalde vaardigheid, invult. Dit impliceert dat een eventueel berekende individuele testscore een grotere meetfout bevat dan wanneer alle items zouden worden afgenomen. De testscore en de onderliggende vaardigheidsscore wijken dus mogelijks sterker van elkaar af dan wanneer een leerling alle items invult. Een manier om de onzekerheid in rekening te brengen die gepaard gaat met

¹⁷ Gebaseerd op von Davier, Gonzalez & Mislevy (2009)

schatten van het onderliggende vaardigheidsniveau, is om meerdere waarden te schatten die de onderliggende verdeling van het werkelijke vaardigheidsniveau van een leerling weerspiegelt. Dit wordt toegepast bij het berekenen van de plausibele waarden.

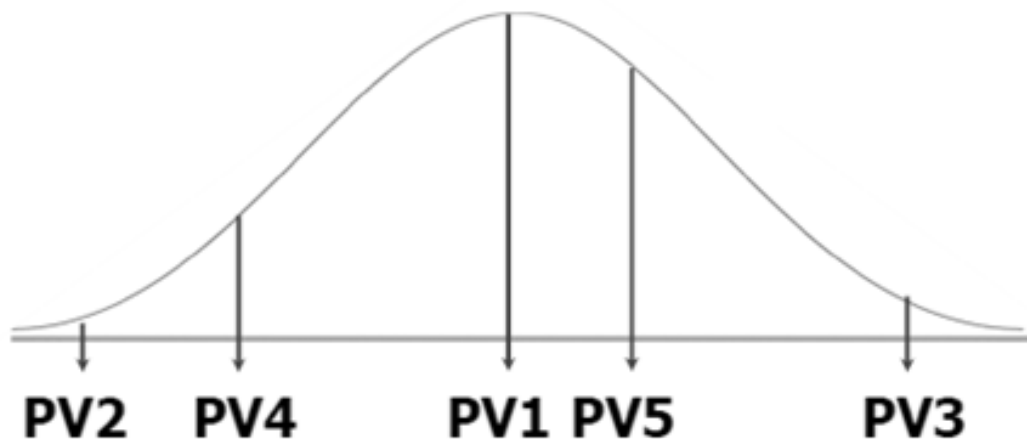
Grafiek 27 toont de veronderstelde verdeling van een vaardigheid (bv. wiskunde) in de populatie. In deze figuur wordt de vaardigheid nagegaan door middel van een toets met 6 items. Het gemiddelde in de populatie is een score van 3. Een toetsscore komt echter niet steeds sluitend overeen met de onderliggende vaardigheidsscore. Deze onderliggende vaardigheidsscore wordt weergegeven door middel van de normaalcurves. Als we bijvoorbeeld naar de curve voor score 2 kijken, is dit de representatie van leerlingen die een score 2 op de toets halen. Hiertoe behoren echter zowel leerlingen die een onderliggende vaardigheidsscore van 2 hebben alsook leerlingen die een score lager of hoger als 2 hebben. Hetzelfde geldt voor de overige normaalcurves.

Grafiek 27. Veronderstelde verdeling in de populatie van een bepaalde vaardigheidsscore, opgesplitst naar toetsscore.



Per leerling die deelneemt aan de toets, worden er vijf waarden, of plausibele waarden, berekend op basis van de veronderstelde vaardigheidsverdeling in de populatie. Grafiek 28 toont hoe plausibele waarden berekend worden van een leerling met een bepaalde toetsscore. Voor deze leerling worden vijf plausibele waarden berekend die zich situeren op de populatieverdeling van leerlingen met deze score. De leerling krijgt op deze manier als het ware vijf onderliggende vaardigheidsscores mee.

Grafiek 28. Illustratie berekening vijf plausible waarden.

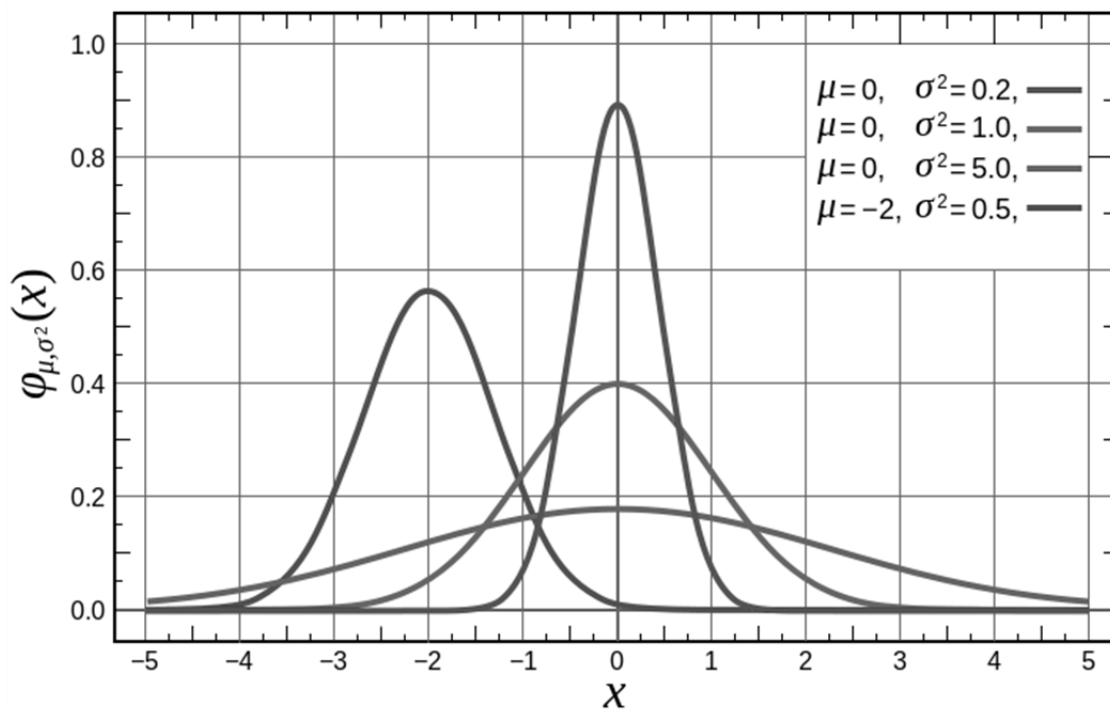


Naast de toetscore van leerlingen, wordt ook informatie over achtergrondkenmerken van leerlingengroepen gebruikt om een meer accurate schatting te geven van de onderliggende vaardigheidsscores van leerlingengroepen. Onder meer hierdoor kunnen plausible waarden niet gebruikt worden als accurate schatters van de individuele vaardigheidsscore, maar kunnen enkel accurate uitspraken over de populatie of over groepen binnen de populatie gedaan worden.

Grafiek 29 toont hoe rekening gehouden wordt met informatie over achtergrondkenmerken van groepen op basis van een hypothetische situatie. De normaalcurves met gemiddelde 0 hebben allen eenzelfde gemiddelde onderliggende vaardigheidsscore, maar een andere verdeling (in Grafiek 27 en Grafiek 28 werd hiervan abstractie gemaakt). Deze verschillende normaalcurven representeren verschillende groepen van leerlingen in de populatie. Veronderstel dat de blauwe curve leerlingen representeert met een hoge SES, de rode curve leerlingen met een middelmatige SES en de gele lijn leerlingen met een lage SES. Deze informatie over de populatieverdeling bij verschillende SES groepen, wordt geïntegreerd in de berekening van de plausible waarden van elke leerling.

Door middel van het opnemen van informatie over achtergrondkenmerken, kan een meer correcte benadering gegeven worden van het onderliggende vaardigheidspatroon van verschillende groepen van leerlingen. Er wordt hierbij rekening gehouden met verscheidene achtergrondkenmerken (zoals geslacht, SES, etniciteit, e.d.) en leerlingen behoren dan ook tot verscheidene groepen. Met al deze informatie over de achtergrondkenmerken van de leerling wordt rekening gehouden bij het berekenen van de vijf uiteindelijke plausible values van elke leerling.

Grafiek 29. Illustratie normaalcurves met verschillend gemiddelde en standaardafwijking.



Analyses waarin prestaties als afhankelijke variabele worden opgenomen, dienen gebruik te maken van de vijf plausible values die per leerling berekend worden door TIMSS. Elke analyse dient herhaald te worden met elke plausible value, waardoor vijf maal dezelfde analyse wordt uitgevoerd, steeds met een andere plausible value als afhankelijke variabele. Het uiteindelijke resultaat is het gemiddelde van deze vijf analyses. Voor een uitvoerige toelichting van het correct gebruik van plausible values in analyses verwijzen we de lezer graag door naar von Davier, Gonzalez en Mislevy (2009).

Wanneer men uitspraken wil doen over de vaardigheidsscores van groepen binnen een populatie, zijn plausible waarden dan ook uitermate geschikt. Onder de aanname dat de veronderstelde populatie- en groepsverdelingen correct zijn, zijn zij namelijk optimale schatters van de vaardigheidsverdeling in de populatie en van verschillende groepen binnen deze populatie.